

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

102540
SEARCH REQUEST FORM

Access DB#

CRFE

Scientific and Technical Information Center

23

Requester's Full Name: ANN LAM Examiner #: 77568 Date: 8/26/03
 Art Unit 1641 Phone Number 306-5560 Serial Number 09/935417
 Mail Box and Bldg/Room Location: 7E12 Results Format Preferred (circle) PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Composite Expandable Device w/ Polymeric Covering + bioactive coating...
 Inventors (please provide full names): Leon V. Rudakov & Mir A. Imrhan
+ Linh Dinh; Ara Davidian & ~~Kevin~~ Kevin T. Larkin
 Earliest Priority Filing Date: Aug 30, 1999

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search peptide that includes the
 amino acid sequence ID No: 1 (claim 19)
 (peptide enhances endothelial cell growth)
 Thank you!
 Ann

RECEIVED
 AUG 27 2003
 INFORMATION DIVISION
 (LIT)

AA

83 1-15

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher _____	NA Sequence (#) _____	STN _____
Searcher Phone # _____	AA Sequence (#) _____	Dialog _____
Searcher Location _____	Structure (#) _____	Questel Orbit _____
Date Searcher Provided _____	Bibliographic _____	Dr. Line _____
Date Submitted _____	Litigation _____	Lexis Nexis _____
Searcher Prep & Review Time _____	Fulltext _____	Sequence Systems _____
Client Prep Time _____	Patent Family _____	WWW/Internet _____
Time _____	Other _____	Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2003, 17:49:44 ; Search time 83 Seconds

28.686 million cell updates/sec

Title: US-09-935-417-1

Sequence: 1 GTPGPQGIAGQRGV 15

Scoring table: BLOSUM62

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
```

Maximum DB seq length: 200000000000

Post-processing:	Minimum Match	0%
------------------	---------------	----

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1980.DAT *
2:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1981.DAT *
3:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1982.DAT *
4:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1983.DAT *
5:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1984.DAT *
6:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1985.DAT *
7:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1986.DAT *
8:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1987.DAT *
9:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1988.DAT *
10:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1989.DAT *
11:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1990.DAT *
12:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1991.DAT *
13:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1992.DAT *
14:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1993.DAT *
15:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1994.DAT *
16:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1995.DAT *
17:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1996.DAT *
18:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1997.DAT *
19:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1998.DAT *
20:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA1999.DAT *
21:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA2000.DAT *
22:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA2001.DAT *
23:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA2002.DAT *
24:	/SDSI1/gcgdata/genseq/genseqp-emb1/AA2003.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	80	100.0	15	12	AAR11114	collagen peptide a
2	80	100.0	15	14	AAR38476	Sequence of peptid
3	80	100.0	15	18	AAW27491	Cell binding pepti
4	80	100.0	15	18	AAW18825	collagen binding p
5	80	100.0	15	20	AAV29591	collagen cell bind
6	80	100.0	15	20	AAV29587	collagen fibronect
7	80	100.0	15	22	AAAG67402	Synthetic peptide
8	80	100.0	15	23	ABP51951	Portion of an al c
9	80	100.0	15	23	ABBI0111	Collagen cell bind

10	80	100.0	16	17	AAR92859	Collagen fragment
11	80	100.0	16	21	AA76868	Collagen receptor
12	80	100.0	19	22	AAB35632	Collagenase cleav
13	80	100.0	25	20	AAV07306	Collagen assembly
14	80	100.0	134	22	ABU53024	Human testes-deri
15	80	100.0	144	22	ABU53008	Human testes-deri
16	80	100.0	144	22	ABU53008	Human testes-deri
17	80	100.0	333	22	AAE02713	Recombinant human
18	80	100.0	333	22	AAB68067	Amino acid sequen
19	80	100.0	416	22	AAE02711	Human alpha (1)
20	80	100.0	416	22	AAB68065	Amino acid sequen
21	80	100.0	500	22	AAE02708	Human alpha (1)
22	80	100.0	500	22	AAB68062	Amino acid sequen
23	80	100.0	510	22	AAE02712	Recombinant human
24	80	100.0	510	22	AAB68066	Amino acid sequen
25	80	100.0	662	22	AAE02718	Human alpha (1)
26	80	100.0	662	22	AAB68072	Amino acid sequen
27	80	100.0	822	22	AAV06240	Mouse recombinant
28	80	100.0	936	22	AAE07107	Gelatin protein.
29	80	100.0	1057	21	AAV84541	Amino acid sequen
30	80	100.0	1057	21	AAV84544	A human collagen
31	80	100.0	1058	21	AAV84403	Amino acid sequen
32	80	100.0	1107	17	AAAR89472	Collagen/decorin(
33	80	100.0	1107	21	AAV84540	Amino acid sequen
34	80	100.0	1169	17	AAAR89469	Collagen/BMP-2B f
35	80	100.0	1169	21	AAV84537	Amino acid sequen
36	80	100.0	1171	17	AAAR89470	Collagen/TGF-beta
37	80	100.0	1171	21	AAV84538	A chimeric collag
38	80	100.0	1341	16	AAV71701	Collagen alpha I
39	80	100.0	1341	21	AAV96122	Collagen type I-a
40	80	100.0	1341	23	ABBB80733	Collagen type I-a
41	80	100.0	1341	23	ABBB09625	Amino acid sequen
42	80	100.0	1341	23	AAE16475	Human collagen
43	80	100.0	1388	17	AAAR89471	Collagen/decorin
44	80	100.0	1388	21	AAV84539	Amino acid sequen
45	80	100.0	1411	21	AAV56800	Human preproalpha

ALIGNMENTS

RESULT 1	
AA011114	
ID	AA011114 standard; peptide; 15 AA

AC AAR1114;

DT	25-MAR-2003	(updated)
DT	17-MAY-1991	(first entry)

DE Collagen peptide analogue.

KW Collagen alpha-1 chain; cell adhesion; vertebrates.

synthetic.

PN W09102537-A.

PD 07-MAR-1991.

PF	13-AUG-1990;	90WO-US04538.
YY		

PR 14-AUG-1989; 89US-0393621.

PA (REGC) UNIV CALIFORNIA.
XY

PI Bhatnagar RS;
XY

DR WPI; 1991-087110/12
XX

PT Synthetic peptide(s) analogous to collagen - promote cell adhesion

PS Claim 1; page 16; 20pp; English

XX This peptide corresponds to a region of the alpha-1 chain of collagen.
 CC It is useful in a compn. for promoting vertebrate cell (esp.
 CC fibroblast) adhesion to a substrate. It is free from natural
 CC folding, glycosylation, cross-linking, hydroxylation and association
 CC with other peptide chains.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 15 AA:

Query Match 100.0%; Score 80; DB 12; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGGRGV 15
 DB 1 GTPGPGIAGGGRGV 15

RESULT 2
 AAR38476 standard; peptide; 15 AA.

XX AAR38476;
 AC
 DT 25-MAR-2003 (updated)
 DT 02-DEC-1993 (first entry)

XX Sequence of peptide P-15 which spans approx. residues 766-780 of the
 DE alpha-1(I) chain of collagen.

XX Synthetic peptide; alpha-1(I) chain; collagen; binding; P-15.

XX Synthetic.

XX W09311781-A1.

XX 24-JUN-1993.

XX 03-DEC-1992; 92MO-US10420.

XX 09-DEC-1991; 91US-0804782.

XX (REGC) UNIV CALIFORNIA.

XX Bhatnagar RS;

XX WPI: 1993-213814/26.

XX Synthetic peptide mimicking collagen binding to cells - used in
 PT composite with bio-material matrix for soft and hard tissue
 PT repair or reconstruction

XX Disclosure; Table 1, page 9; 26pp; English.

XX The P-15 peptide spans approx. residues 766-780 of the alpha-1(I)
 CC chain of collagen. The P-15 region does not occur as a natural
 CC fragment of collagen nor is it a product of natural enzymatic
 CC cleavage. The P-15 region represent half of one turn of the collagen
 CC triple helix. The sequence contd. in P-15 can acquire a conformation
 CC dramatically different from the triple helical conformation
 CC generally observed in the rest of the collagen molecule. AAR38477-82
 CC is a family of synthetic peptide fragments of P-15. They mimic the
 CC cell binding domain of collagen. The domain includes a core
 CC sequence that, at physiologic conditions, is folded in a beta-bend
 CC formed at the Ile-Ala.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 15 AA:

Query Match 100.0%; Score 80; DB 14; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGGRGV 15
 DB 1 GTPGPGIAGGGRGV 15

RESULT 3
 AAM27491 standard; peptide; 15 AA.

XX AAM27491;

XX 20-APR-1998 (first entry)

XX Cell binding peptide #1 derived from collagen.

XX Bioreactor; packing material; cell culture; collagen alpha1(I) chain;
 KW cell binding peptide; matrix.

XX Synthetic.
 OS Mammalia.

XX US5674848-A.

XX 07-OCT-1997.

XX 03-AUG-1994; 94US-0285570.

XX 14-AUG-1989; 89US-0393621.
 PR 09-DEC-1991; 91US-0804782.

XX (REGC) UNIV CALIFORNIA.

XX Bhatnagar RS;

XX WPI: 1997-502373/46.

XX Bioreactor packing material for cell culture - comprising matrix
 PT coated with cell binding peptide

XX Claim 1; Col 18; 13pp; English.

XX The present peptide sequence corresponds to a region of the alpha1(I)
 CC chain of collagen which is sometimes referred to as "P-15". It can be
 CC used as a cell binding peptide in a new packing material, which is useful
 CC for cell culture in a bioreactor. The material comprises a matrix formed
 CC of a biomaterial, i.e. a material that is biologically compatible for in
 CC vivo applications and for cell culture in vitro, and the cell binding
 CC peptide. A bioreactor containing the packing material can be used to
 CC culture cells, e.g. mammalian cells for the production of monoclonal
 CC antibodies. The peptides are more effective than collagen in promoting
 CC cell attachment.

XX Sequence 15 AA:

Query Match 100.0%; Score 80; DB 18; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGGRGV 15
 DB 1 GTPGPGIAGGGRGV 15

RESULT 4
 AAM18825 standard; peptide; 15 AA.

XX AAM18825;

XX 25-MAR-2003 (updated)
 DT 05-JAN-1998 (first entry)

DE Collagen binding peptide mimic 1.
XX
KW implant; biomaterial matrix; enhanced cell binding; collagen;
KW beta-bend; fold; substrate; reconstructive surgery; bone; ligament;
KW repair; tooth.
XX
OS Synthetic.
XX
XX US635482-A ★
XX 03-JUN-1997.
XX
XX 22-JUL-1994; 94US-0278878.
XX PF
PR 22-JUL-1994; 94US-0278878.
PR 14-AUG-1989; 89US-0393621.
PR 09-DEC-1991; 91US-0804782.
XX
PA (REGC) UNIV CALIFORNIA.
XX
XX Bhatnagar RS;
XX WPI; 1997-309859/28.
XX
PT Implant bearing cell-binding collagen-mimetic peptide - for
PT promoting cell attachment
XX
XX Claim 1; Column 18; 12pp; English.
XX
XX New implants comprise a biomaterial matrix and a peptide carried by the
CC matrix where the peptide has enhanced cell binding with respect to
CC collagen and has a domain that mimics collagen binding to cells, the
CC domain including at least -Ile-Ala-folded in a beta-bend at
CC physiological conditions. The peptide is one of AAM18825-34 or one of 3
CC tripeptides (Nac-Ile-Ala-Ala; Ile-Ala-beta Ala; and Nac-Ile-Ala-N-Me).
CC The implant is used as a substrate for growing cells, e.g. for use in
CC reconstructive surgery, e.g. for bone or ligament repair or as tooth
CC implants. The peptide promotes cell attachment to the matrix and also
CC cell migration into the matrix when the matrix is porous.
CC (updated on 25-MAR-2003 to correct pf field.)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 80; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPGPGGIAGORGVV 15
DB 1 GTPGPGGIAGORGVV 15
|||||

RESULT 5
AAAY29991
ID AAY29991 standard; peptide: 15 AA.
XX
XX AAY29991;
XX AC
XX 02-DEC-1999 (first entry)
XX DT
XX DE Collagen cell binding domain mimotope #1.
KW Collagen; cell binding domain; biomaterial; soft tissue repair;
KW hard tissue repair; reconstruction; cell surface receptor;
KW fibronectin; beta-bend; cartilage; tendon; ligament; bone.
XX
XX Synthetic.
XX OS
XX US5958428-A ★
XX PN
PD 28-SEP-1999.
XX PD
XX 20-MAY-1997; 97US-0859610.

22-JUL-1994; 94US-0278878.
14-AUG-1989; 89US-0393621.
09-DEC-1991; 91US-0804782.
(REGC) UNIV CALIFORNIA.
Bhatnagar RS;
WPI; 1999-561009/47.
Synthetic peptide additives with enhanced collagen binding affinities useful for the production of apparatus for soft tissue, cartilage and bone repair -
Claim 3; Column 25; 16pp; English.
The present invention describes synthetic peptide additives (SPAs) with enhanced collagen binding affinities. AAY29587 to AAY30000 represent specifically claimed examples of the SPAs. The additives comprise domains that mimic the binding sites of collagen to cells (but with higher affinity) and promote cell attachment when the additives are carried on repair or reconstructive apparatus. The SPA may be used in the construction of apparatus for soft tissue, cartilage, tendon, ligament and bone repair. The SPA mimics and enhances the binding of cells to the tissue repair apparatus
Sequence 15 AA;
Query Match 100.0%; Score 80; DB 20; Length 15;
Best Local Similarity 100.0%; Pred. NO. 6.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTPGPGIAGQGRGV 15
|||||
DB 1 GTPGPGIAGQGRGV 15
RESULT 6
AAY29587
AAY29587 standard; peptide; 15 AA.
XX
AC AAY29587;
XX
DX 18-OCT-1999 (first entry)
XX
DE Collagen fibronectin binding region oligopeptide.
XX
KW Collagen; fibronectin binding region; tissue regeneration; implant;
XX internal wound site; biodegradable microparticle.
XX
OS Unidentified.
XX
PN WO933447-A2.
XX
PD 08-JUL-1999.
XX
PF 24-DEC-1998; 98WO-US27596.
XX
PR 30-DEC-1997; 97US-0000638.
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Yannas IV;
XX
DR WPI; 1999-493795/41.
XX
PT Biodegradable microparticles for tissue regeneration at an internal
XX wound site
XX
PS Disclosure; Page 8; 25pp; English.
XX
CC The present invention describes a porous biodegradable microparticle (I)

CC for tissue regeneration at an internal wound site in a subject. The
 CC pores of (I) have a diameter 1-300 nm m; (I) has a minimum water content
 CC of at least about 80%, a minimum specific surface area of at least about
 CC 103 mm² per cm³ and a diameter 10-1000 micro m. between about 20-80% by
 CC weight of (I) is biodegraded at the wound site during the time period
 CC required for a wound of about the same severity, size and tissue type to
 CC complete about one half of the contraction which normally takes place in
 CC the absence of (I); and (II) comprises: (i) a three dimensional network
 CC of polymers which is substantially insoluble under physiological
 CC conditions; and (ii) one or more specific cell-binding fragments.
 CC Methods using (I) may be used to treat internal injuries caused to
 CC internal organs by disease or trauma, and to inhibit wound contraction
 CC and scar formation. The methods work by preventing contractile cells in
 CC the vicinity of a wound site (accidentally or surgically induced) on an
 CC internal organ from inducing contraction at the lesion site. The tissue
 CC regeneration methods greatly improve the clinical outcomes of patients
 CC with internal organ and tissue injuries. The present sequence represents
 CC a collagen fibronectin binding region oligopeptide which is used in as
 CC part of an example of a specific cell binding fraction which is included
 CC in a 3-dimensional network of the regeneration template from the present
 CC invention.

XX Sequence 15 AA;

Query Match 100.0%; Score 80; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPPGGAGGRCVV 15
 |||||
 Db 1 GTPPGGAGGRCVV 15

RESULT 8
 AAG67402
 ID AAG67402 standard; peptide: 15 AA.

XX AAG67402;

DT 13-NOV-2001 (first entry)

XX Synthetic peptide mimicking cell binding domain of collagen.

XX Cell binding; collagen; cell migration; collagen receptor; tissue repair;
 KW metalloproteinase; prolyl hydroxylase; tissue reconstruction; arthritis;
 KW bone repair; tooth implant; ligament repair; scar tissue; osteoporosis;
 KW bone disease; cartilage repair; joint disease; tendon repair.

XX Synthetic.

XX US6268348-B1.

XX 31-JUL-2001.

XX 08-JUN-1999; 99US-0328347.

XX 22-JUL-1994; 94US-0278878.

XX 20-MAY-1997; 97US-0859610.

XX 14-AUG-1989; 89US-0393621.

XX 09-DEC-1991; 91US-0804782.

XX (REGC) UNIV CALIFORNIA.

XX Bhattachar RS;

XX WPI; 2001-540321/60.

XX New collagen binding synthetic peptide useful for soft and hard tissue
 PT repair e.g. bone repairs comprises a family of amino acid sequence -
 XX Claim 1; Column 25; 16pp; English.

XX The present sequence represents a synthetic peptide, which mimics the

CC cell binding domain of collagen. The cell binding ability of the
 CC peptide is enhanced with respect to collagen. The peptide promotes cell
 CC migration into porous lattices; binds to collagen receptors; induces
 CC metalloproteinases; can down regulate prolyl hydroxylase and collagen;
 CC inhibits cell binding to collagen or inhibits cell migration in vitro.
 CC The peptide is used for soft and hard tissue repair or reconstruction;
 CC e.g. bone repair, tooth implants and ligament repair for in vitro uses;
 CC as an inhibitor of collagen synthesis to block formation of scar tissue
 CC and thus promotes scarless healing; as bone filling/fusion for
 CC osteoporosis and other bone diseases; cartilage repair for arthritis and
 CC other joint disease and tendon repair; for soft tissue repair e.g. nerve,
 CC organ, skin, vascular, muscle and ophthalmic applications.

XX Sequence 15 AA;

Query Match 100.0%; Score 80; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPPGGAGGRCVV 15
 |||||
 Db 1 GTPPGGAGGRCVV 15

RESULT 8
 ABP51951
 ID ABP51951 standard; peptide: 15 AA.

XX ABP51951;

DT 08-OCT-2002 (first entry)

XX Portion of an al chain or collagen peptide sequence SEQ ID NO:1.

XX Delivery; blood; collagen; occlusion; blood vessel; saphenous vein graft.

XX Synthetic.

XX US2002062145-A1.

XX 23-MAY-2002.

XX 22-AUG-2001; 2001US-0935417.

XX 30-AUG-1999; 99US-0385691.

XX (CARD-) CARDIOVASC INC.

XX Rudakov, LV Imran MA, Dinh L, Davidian A, Larkin KT;

XX WPI; 2002-582305/62.

XX Composite expandable device for treating occlusions in blood vessels,
 PT e.g., saphenous vein grafts, comprises polymeric covering and bioactive
 PT coating

XX Disclosure; Page 4; 10pp; English.

XX The present invention describes a composite expandable device for
 CC delivery into a blood vessel comprising an expandable support frame,
 CC an impervious polymer sleeve extending over the support frame, and a
 CC coating disposed on the inner and outer surfaces of the polymer sleeve
 CC for enhancing endothelial cell growth on the polymer sleeve. Also
 CC described is delivery apparatus for an expandable device comprising a
 CC shaft, and a balloon mounted on the shaft, where the shaft has a lumen
 CC for inflating and deflating the balloon. The balloon is formed with
 CC intermediate portions adapted to receive the expandable device, and
 CC radiopaque markers carried within the balloon and sized to prevent from
 CC being dislodged during deployment by the delivery apparatus. The
 CC apparatus is used for treating occlusions or partial occlusions in
 CC blood vessels, particularly saphenous vein grafts. The present
 CC sequence represents a portion of an al chain of collagen, which is
 CC given in the exemplification of the present invention.

XX Sequence 15 AA:

Query Match 100.0%; Score 80; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPPGGIAGRGVY 15
DB 1 GTPPGGIAGRGVY 15

RESULT 9

AB10111 standard; peptide: 15 AA.

AC ABB10111;

DT 12-JUL-2002 (first entry)

DE Collagen cell binding domain mimic peptide P-15.

KW Collagen; bone; repair; bone graft; tissue engineering; fibroblast;
radiation therapy; bone damage.

OS Synthetic.

PN WO200182773-A2.

PD 08-NOV-2001.

PF 29-MAR-2001; 2001WO-US10404.

PR 28-APR-2000; 2000US-0561554.

PA (REGC) UNIV CALIFORNIA.

PI Bhatnagar RS, Qian JY;

DR WPI: 2002-034479/04.

PT Preparation of bone repair apparatus comprises seeding at least some of
cultured tissue cells on biologically compatible structure having
collagen mimic and incubating seeded cells under cell growth conditions

PS Claim 7; Page 6; 23pp; English.

XX The invention relates to a bone repair apparatus that is prepared by
growing harvested fibroblasts under cell growth conditions to form
cultured tissue cells, seeding at least some of the cultured tissue
cells on a biologically compatible structure having a collagen mimic, and
incubating the seeded cells under cell growth conditions, where the
seeded cells differentiate into an osteogenic phenotype. Methods of the
invention are useful for preparing bone repair apparatus for use as a
bone graft. The fibroblast cells from the recipient can be easily
harvested with minimal invasion and trauma to the patient. By contrast to
other methods, the fibroblast is plentiful and easily obtained with
minimal trauma and the inventive method is able to obtain living bone
-like cells and they, together with the biologically compatible
structure, yield a tissue engineered bone graft. This can integrate with
host bone when implanted in the patient, and repopulates host sites
lacking viable cells because of disease or radiation therapy. The current
sequence represents a collagen cell binding domain mimic peptide P-15.
This 15 amino acid peptide has the same sequence as a particular, small
region in the alpha(1) chain of collagen.

SO Sequence 15 AA:

Query Match 100.0%; Score 80; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPPGGIAGRGVY 15
DB 1 GTPPGGIAGRGVY 15

RESULT 10

AA92859 standard; peptide: 16 AA.

AC AA92859;

DT 03-OCT-1996 (first entry)

DE Collagen fragment P-15 as positive control for cell adhesion.

KW Inter cellular adhesion; stimulation; inhibition; skin graft;
synthetic blood vessel; coating; endothelial cell; epidermal cell;
chemotactic attractor; wound healing; organ transplantation;
thrombosis; arteriosclerosis; cancer metastases.

OS Synthetic.

PN DE4430601-A1.

PD 29-FEB-1996.

PF 22-AUG-1994; 94DE-4430601.

PR 22-AUG-1994; 94DE-4430601.

PA (BEIE) BEIERSDORF AG.

PI Doerschner A, Eichner W, Kock K, Mielke H;

DR WPI: 1996-130242/14.

PT Peptide(s) that stimulate or inhibit cell to cell adhesion - used
e.g. to coat synthetic blood vessels with endothelial cells, to
prepare, or increase growth of skin grafts, to prevent thrombosis
etc.

PS Example 1; Page 7; 18pp; German.

XX Peptides contg. the highly generic sequence AA5-AA4-AA3-AA2-AA1-(AAx)n
where AA5 is Glu, Ser, Asp or Asn; AA4 is Leu or Ser, AA3 is Leu, Ile,
Phe or Gly; AA2 is Asp, Leu, Asn or Ser; AA1 is Gly, Pro or Asp; AAx
is any amino acid and n = 0 or 1 are claimed; AA5 or AA5-AA4 may be
absent. When two or more such peptides are attached to a carrier, the
product can be used for stimulating adhesion of eukaryotic cells in
vitro. Particular applications include coating synthetic blood vessels
with endothelial cells, preparing skin grafts using epithelial cells
or stimulating wound healing. When a single peptide is used it may
inhibit intercellular adhesion, making it useful for preventing
thrombosis or arteriosclerosis or to suppress cancer metastases. The
peptides can also be used as chemotactic attractors and for detecting/
quantifying cell-cell adhesion in vitro.
CC The present sequence is a fragment of the alpha-1 chain of collagen
which was used as a positive control in a cell adhesion assay on the
novel peptides.

SO Sequence 16 AA:

Query Match 100.0%; Score 80; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Collagen type II
US-08-931-820-3

Query Match 88.8%; Score 71; DB 3; Length 1060;
Best Local Similarity 80.0%; Pred. No. 0.017;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTPGPGIAGORGIV 15
| | | | | : | | | | | : |
Db 788 GTPGPGIAGORGIV 802

RESULT 14
US-08-963-825-20
Sequence 20, Application US/08963825
Patent No. 6110689

GENERAL INFORMATION:
APPLICANT: Qvist, Per
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
TITLE OF INVENTION: Disorders Associated with the Metabolism of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,825
FILING DATE:

CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/187,319
FILING DATE: 21-JAN-1994

ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: COLLAGEN -ALPHA 1 (II)
US-08-963-825-20

Query Match 88.8%; Score 71; DB 3; Length 1418;
Best Local Similarity 80.0%; Pred. No. 0.023;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTPGPGIAGORGIV 15
| | | | | : | | | | | : |
Db 900 GTPGPGIAGORGIV 914

RESULT 15
US-09-010-999-1
Sequence 1, Application US/09010999
Patent No. 6132976

GENERAL INFORMATION:
APPLICANT: Poole, Anthony R.
APPLICANT: Hollander, Anthony P.
APPLICANT: Billingham, R. C.
TITLE OF INVENTION: IMMUNOASSAYS FOR THE MEASUREMENT OF
TITLE OF INVENTION: COLLAGEN DENATURATION AND CLEAVAGE IN CARTILAGE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,999
FILING DATE: 22-JAN-1998
CLASSIFICATION: 4335

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/448,501
FILING DATE: 17-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,123
FILING DATE: 04-DEC-1992

ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 032931/0212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1418 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Human Type II Collagen
US-09-010-999-1

Query Match 88.8%; Score 71; DB 3; Length 1418;
Best Local Similarity 80.0%; Pred. No. 0.023;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTPGPGIAGORGIV 15
| | | | | : | | | | | : |
Db 900 GTPGPGIAGORGIV 914

Search completed: August 29, 2003, 18:30:11
Job time : 30 secs

;; TITLE OF INVENTION: Disorders Associated with the Metabolism of
;; NUMBER OF SEQUENCES: 21
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Darby & Darby PC
;; STREET: 805 Third Avenue
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10022
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/548,608
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/187,319
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gogoris, Adda C
;; REGISTRATION NUMBER: 29,714
;; REFERENCE/DOCKET NUMBER: 4305/08701
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-527-7700
;; TELEFAX: 212-753-6237
;; TELEX: 236687
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1341 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; IMMEDIATE SOURCE:
;; CLONE: COLLAGEN ALPHA 1 (I)
;; US-09-548-608-18

Query Match 100.0%; Score 80; DB 4; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGRGV 15
DB 823 GTPGPGIAGRGV 837

RESULT 11
US-09-585-887-9
; Sequence 9, Application US/09585887
; Patent No. 6413742
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R
; APPLICANT: Chang, Robert
; APPLICANT: McMullin, Hugh
; APPLICANT: Hitzeman, Ronald A.
; APPLICANT: Chisholm, George
; TITLE OF INVENTION: NOVEL METHODS FOR THE PRODUCTION OF GELATIN AND
; TITLE OF INVENTION: FULL-LENGTH TRIPLE HELICAL COLLAGEN IN RECOMBINANT
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 225002030400
; CURRENT APPLICATION NUMBER: US/09/585,887
; CURRENT FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/289,578
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/084,828
; PRIOR FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9

;; LENGTH: 1461
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; US-09-585-887-9

Query Match 100.0%; Score 80; DB 4; Length 1461;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGRGV 15
DB 944 GTPGPGIAGRGV 958

RESULT 12
US-09-289-578-9
; Sequence 9, Application US/09289578
; Patent No. 6428978
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R
; APPLICANT: Chang, Robert
; APPLICANT: McMullin, Hugh
; APPLICANT: Hitzeman, Ronald A.
; APPLICANT: Chisholm, George
; TITLE OF INVENTION: NOVEL METHODS FOR THE PRODUCTION OF GELATIN AND
; TITLE OF INVENTION: FULL-LENGTH TRIPLE HELICAL COLLAGEN IN RECOMBINANT
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 225002030400
; CURRENT APPLICATION NUMBER: US/09/289,578
; CURRENT FILING DATE: 1999-04-10
; PRIOR APPLICATION NUMBER: 60/084,828
; PRIOR FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1461
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-289-578-9

Query Match 100.0%; Score 80; DB 4; Length 1461;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGRGV 15
DB 944 GTPGPGIAGRGV 958

RESULT 13
US-08-931-820-3
; Sequence 3, Application US/08931820
; Patent No. 6010863
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Assay for collagen degradation
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,820
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 96202596.1
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1060 amino acids
; TYPE: amino acid

MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: COLLAGEN ALPHA 1 (1)
US-08-963-825-18

Query Match 100.0%; Score 80; DB 3; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGORGV 15
|||||
DB 823 GTPGPGIAGORGV 837

RESULT 8
US-09-500-811-18
Sequence 18, Application US/09500811
Patent No. 6323314
GENERAL INFORMATION:
APPLICANT: Qvist, Per
APPLICANT: Bonde, Martin
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
TITLE OF INVENTION: Disorders Associated with the Metabolism of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,811
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/187,319
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1341 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: COLLAGEN ALPHA 1 (1)
US-09-500-811-18

Query Match 100.0%; Score 80; DB 4; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGORGV 15
|||||

DB 823 GTPGPGIAGORGV 837

RESULT 9
US-09-570-573-18
Sequence 18, Application US/09570573
Patent No. 6342361
GENERAL INFORMATION:
APPLICANT: Qvist, Per
APPLICANT: Bonde, Martin
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
TITLE OF INVENTION: Disorders Associated with the Metabolism of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/570,573
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/187,319
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1341 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: COLLAGEN ALPHA 1 (1)
US-09-570-573-18

Query Match 100.0%; Score 80; DB 4; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGORGV 15
|||||

DB 823 GTPGPGIAGORGV 837

RESULT 10
US-09-548-608-18
Sequence 18, Application US/09548608
Patent No. 6355442
GENERAL INFORMATION:
APPLICANT: Qvist, Per
APPLICANT: Bonde, Martin
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of

EARLIER FILING DATE: 1998-09-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO: 3
LENGTH: 25
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: COL1A1 Binding Peptides
US-09-517-866-3

Query Match 100.0%; Score 80; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPGPGIAGRGV 15
|||||
DB 9 GTPGPGIAGRGV 23

RESULT 5
US-09-219-849-49
Sequence 49, Application US/09219849
Patent No. 6150081
GENERAL INFORMATION:
APPLICANT: VAN HEERDE, GEORGE V.
APPLICANT: VAN RIJN, ALEXIS C.
APPLICANT: BOUWSTRA, JAN B.
APPLICANT: DE WOLF, FREDERIK A.
APPLICANT: MOORROEK, ANDREAS
APPLICANT: WERTEN, MARC W.T.
APPLICANT: WIND, RICHELIE D.
APPLICANT: VAN DEN BOSCH, TANJA J.
TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
FILE REFERENCE: 2728-2
CURRENT APPLICATION NUMBER: US/09/219,849
CURRENT FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 49
LENGTH: 822
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-219-849-49

Query Match 100.0%; Score 80; DB 3; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.00069;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPGPGIAGRGV 15
|||||
DB 768 GTPGPGIAGRGV 782

RESULT 6
US-08-931-820-1
Sequence 1, Application US/08931820
Patent No. 6010863
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Assay for collagen degradation
NUMBER OF SEQUENCES: 4
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/931,820
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 96202596.1
FILING DATE:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1057 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Collagen type I
US-08-931-820-1

Query Match 100.0%; Score 80; DB 3; Length 1057;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPGPGIAGRGV 15
|||||
DB 786 GTPGPGIAGRGV 800

RESULT 7
US-08-963-825-18
Sequence 18, Application US/08963825
Patent No. 6110689
GENERAL INFORMATION:
APPLICANT: Qvist, Per
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
TITLE OF INVENTION: Disorders Associated with the Metabolism of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRES:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,825
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/187,319
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adida C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1341 amino acids
TYPE: amino acid
TOPOLOGY: linear

Query Match 100.0%; Score 80; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGPGGIAGRGVY 15
|||||
DB 1 GTGPGGIAGRGVY 15

RESULT 2

US-09-328-347A-1
; Sequence 1, Application US/09328347A
; Patent No. 6268348
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Rajendra S.
; TITLE OF INVENTION: SYNTHETIC COMPOUNDS AND COMPOSITIONS
; TITLE OF INVENTION: WITH ENHANCED CELL BINDING
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Hsue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111-4106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,347A
; FILING DATE: 08-JUN-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/859,610
; FILING DATE: 20-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,878
; FILING DATE: 22-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/804,782
; FILING DATE: 09-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/393,621
; FILING DATE: 14-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Siebert, J. Suzanne
; REGISTRATION NUMBER: 28,758
; REFERENCE/DOCKET NUMBER: 2500,066US5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-248-5500
; TELEFAX: 415-362-5418.
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-09-328-347A-1

Query Match 100.0%; Score 80; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGPGGIAGRGVY 15
|||||
DB 1 GTGPGGIAGRGVY 15

RESULT 3

US-09-010-999-9
; Sequence 9, Application US/09010999
; Patent No. 6132976
; GENERAL INFORMATION:
; APPLICANT: Poole, Anthony R.
; APPLICANT: Hollander, Anthony P.
; APPLICANT: Billingham, R. C.
; TITLE OF INVENTION: IMMUNOASSAYS FOR THE MEASUREMENT OF
; TITLE OF INVENTION: COLLAGEN DENATURATION AND CLEAVAGE IN CARTILAGE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,999
; FILING DATE: 22-JAN-1998
; CLASSIFICATION: 4335
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/448,501
; FILING DATE: 17-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,123
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 032931/0212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELER: 904136
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-010-999-9

Query Match 100.0%; Score 80; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGPGGIAGRGVY 15
|||||
DB 4 GTGPGGIAGRGVY 18

RESULT 4

US-09-517-866-3
; Sequence 3, Application US/09517866
; Patent No. 6472504
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, Darwin J.
; APPLICANT: FERTALA, Andrez J.
; TITLE OF INVENTION: INHIBITORS OF COLLAGEN ASSEMBLY
; FILE REFERENCE: 209598,0111/2801
; CURRENT APPLICATION NUMBER: US/09/517,866
; EARLIER FILING DATE: 2000-03-03
; EARLIER APPLICATION NUMBER: 60/058,353
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: PCT/US98/18838

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2003, 18:25:20 ; Search time 29 seconds

(without alignments)
21.885 Million cell updates/sec

Title: US-09-935-417-1

Perfect score: 80

Sequence: 1 GTPGPGIAGRGVY 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*

2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*

3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*

4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*

5: /cgn2_6/prodata/1/1aa/6CTUS.COMB.pep:*

6: /cgn2_6/prodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	15	2	US-08-859-610A-1
2	80	100.0	15	3	US-09-328-347A-1
3	80	100.0	19	3	US-09-010-999-9
4	80	100.0	25	4	US-09-517-866-3
5	80	100.0	822	3	US-09-219-849-49
6	80	100.0	1057	3	US-08-931-820-1
7	80	100.0	1341	3	US-08-963-825-18
8	80	100.0	1341	4	US-09-500-811-18
9	80	100.0	1341	4	US-09-570-573-18
10	80	100.0	1341	4	US-09-548-608-18
11	80	100.0	1461	4	US-09-585-578-9
12	80	100.0	1461	4	US-09-289-578-9
13	80	100.0	1060	3	US-08-931-820-3
14	80	100.0	1418	3	US-08-963-825-20
15	80	100.0	1418	3	US-09-010-999-1
16	80	100.0	1418	4	US-09-500-811-20
17	80	100.0	1418	4	US-09-570-573-20
18	80	100.0	1418	4	US-09-548-608-20
19	80	100.0	1418	4	US-09-548-608-20
20	80	100.0	1418	4	US-09-548-608-20
21	80	100.0	1418	4	US-09-548-608-20
22	80	100.0	1418	4	US-09-548-608-20
23	80	100.0	1418	4	US-09-548-608-20
24	80	100.0	1418	4	US-09-548-608-20
25	80	100.0	1418	4	US-09-548-608-20
26	80	100.0	1418	4	US-09-548-608-20
27	80	100.0	1418	4	US-09-548-608-20

28	56	70.0	19	3	US-09-010-999-10	Sequence 10, Appl
29	56	70.0	1024	3	US-08-931-820-2	Sequence 2, Appl
30	56	70.0	1366	3	US-08-963-825-19	Sequence 19, Appl
31	56	70.0	1366	4	US-09-500-811-19	Sequence 19, Appl
32	56	70.0	1366	4	US-09-570-573-19	Sequence 19, Appl
33	56	70.0	1366	4	US-09-548-608-19	Sequence 19, Appl
34	56	70.0	1366	4	US-09-585-887-19	Sequence 10, Appl
35	56	70.0	1366	4	US-09-585-887-19	Sequence 10, Appl
36	54	67.5	228	3	US-09-219-849-38	Sequence 38, Appl
37	54	67.5	279	3	US-09-010-999-2	Sequence 2, Appl
38	54	67.5	417	1	US-08-175-155-69	Sequence 69, Appl
39	54	67.5	417	1	US-08-477-509B-104	Sequence 104, App
40	54	67.5	417	1	US-08-642-255-102	Sequence 102, App
41	54	67.5	417	2	US-08-707-237A-76	Sequence 76, Appl
42	54	67.5	417	2	US-08-482-085B-104	Sequence 104, App
43	54	67.5	417	4	US-09-444-791A-104	Sequence 104, App
44	54	67.5	829	1	US-08-642-255-132	Sequence 132, App
45	54	67.5	829	1	US-08-397-633A-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-08-859-610A-1
; Sequence 1, Application US/08859610A
; Patent No. 5958428
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Rajendra S.
; TITLE OF INVENTION: SYNTHETIC COMPOUNDS AND COMPOSITIONS
; TITLE OF INVENTION: WITH ENHANCED CELL BINDING
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Hae P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111-4106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,610A
; FILING DATE: 20-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,878
; FILING DATE: 22-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/804,782
; FILING DATE: 09-DEC-1991
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/393,621
; FILING DATE: 14-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Siebert, J. Suzanne
; REGISTRATION NUMBER: 28,758
; REFERENCE/DOCKET NUMBER: 2500.066US4
; TELEPHONE: 415-248-5500
; TELEFAX: 415-362-5418
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-859-610A-1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2003, 18:26:40 ; Search time 25 Seconds
(without alignments)
82.069 Million cell updates/sec

Title: US-09-935-417-1

Sequence: 1 GTPGPGIAGRGV 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 510680 seqs, 136781880 residues

Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PC1US_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	15	9 US-09-113-696B-17	Sequence 17, Appl
2	80	100.0	15	9 US-09-816-737-1	Sequence 1, Appl
3	80	100.0	15	9 US-09-935-417-1	Sequence 1, Appl
4	80	100.0	15	15 US-10-133-289-1	Sequence 1, Appl
5	80	100.0	15	15 US-10-176-401-1	Sequence 1, Appl
6	80	100.0	15	15 US-10-017-193-1	Sequence 1, Appl
7	80	100.0	25	15 US-10-279-991-3	Sequence 3, Appl
8	80	100.0	1341	15 US-10-058-124-18	Sequence 18, Appl
9	80	100.0	1464	12 US-10-301-822-38	Sequence 28, Appl
10	80	100.0	1464	15 US-10-060-036-159	Sequence 159, App
11	80	100.0	1464	15 US-10-171-311-36	Sequence 36, Appl
12	80	100.0	1464	15 US-10-216-705-21	Sequence 21, Appl
13	80	100.0	1464	15 US-10-149-352-2	Sequence 65, Appl
14	80	100.0	1464	15 US-10-177-293-65	Sequence 1, Appl
15	71	88.8	234	11 US-09-895-674-1	Sequence 1, Appl

16	71	88.8	714	15	US-10-233-885-44	Sequence 44, Appl
17	71	88.8	714	15	US-10-231-581-44	Sequence 44, Appl
18	71	88.8	1014	12	US-10-194-441A-1	Sequence 1, Appl
19	71	88.8	1014	12	US-10-194-441A-48	Sequence 48, Appl
20	71	88.8	1418	15	US-10-058-124-20	Sequence 20, Appl
21	60	75.0	459	9	US-09-789-561-97	Sequence 97, Appl
22	58	72.5	1496	12	US-10-301-822-35	Sequence 35, Appl
23	58	72.5	1496	15	US-10-177-293-70	Sequence 70, Appl
24	57	71.2	1078	15	US-10-058-124-21	Sequence 21, Appl
25	57	71.2	1466	12	US-10-301-822-33	Sequence 33, Appl
26	57	71.2	1466	15	US-10-177-293-68	Sequence 68, Appl
27	56	70.0	674	9	US-09-925-299-979	Sequence 979, App
28	56	70.0	674	11	US-09-925-299-979	Sequence 979, App
29	56	70.0	1366	12	US-10-301-822-31	Sequence 31, Appl
30	56	70.0	1366	15	US-10-171-311-38	Sequence 38, Appl
31	56	70.0	1366	15	US-10-058-124-19	Sequence 19, Appl
32	54	67.5	33	9	US-09-864-761-39163	Sequence 39163, A
33	54	67.5	417	15	US-10-096-986-104	Sequence 104, App
34	54	67.5	837	15	US-10-096-986-103	Sequence 103, App
35	53	66.2	61	15	US-10-096-986-102	Sequence 102, App
36	53	66.2	680	15	US-10-177-293-59	Sequence 59, Appl
37	53	66.2	742	15	US-10-203-860-4	Sequence 4, Appl
38	52	65.0	403	9	US-09-925-302-689	Sequence 689, App
39	52	65.0	520	10	US-09-978-295A-614	Sequence 614, App
40	52	65.0	520	10	US-09-978-697-614	Sequence 614, App
41	52	65.0	520	10	US-09-978-192A-614	Sequence 614, App
42	52	65.0	520	10	US-09-999-832A-614	Sequence 614, App
43	52	65.0	520	11	US-09-978-189-614	Sequence 614, App
44	52	65.0	520	11	US-09-978-608A-614	Sequence 614, App
45	52	65.0	520	11	US-09-978-585A-614	Sequence 614, App

ALIGNMENTS

RESULT 1
US-09-113-696B-17
Sequence 17, Application US/09113696B
Patent No. US2002010134A1
GENERAL INFORMATION:
APPLICANT: Bhatnagar, Rajendra S.
APPLICANT: Qian, Jing Jing
TITLE OF INVENTION: PEPTIDE COMPOSITIONS MIMICKING TGF-BETA
FILE REFERENCE: 6510-215CIP2
CURRENT APPLICATION NUMBER: US/09/113,696B
CURRENT FILING DATE: 1998-07-10
PRIOR APPLICATION NUMBER: 08/742,256
PRIOR FILING DATE: 1996-10-31
PRIOR APPLICATION NUMBER: 08/431,954
PRIOR FILING DATE: 1995-05-01
NUMBER OF SEQ ID NOS: 42
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: TGF-beta mimic
US-09-113-696B-17
Query Match 100.0%, Score 80, DB 9, Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTPGPGIAGRGV 15
|||||
DB 1 GTPGPGIAGRGV 15
RESULT 2
US-09-816-737-1

```
Sequence 1, Application US/09616737
Patent No. US20020037853A1
GENERAL INFORMATION:
APPLICANT: Bhatnagar, Rajendra S.
TITLE OF INVENTION: "Synthetic Compounds and Compositions
FILE REFERENCE: 06510223CON2
CURRENT APPLICATION NUMBER: US/09/816,737
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: 09/328,347
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 08/859,610
PRIOR FILING DATE: 1997-05-20
PRIOR APPLICATION NUMBER: 08/278,878
PRIOR FILING DATE: 1994-07-22
PRIOR APPLICATION NUMBER: 07/804,782
PRIOR FILING DATE: 1991-12-09
PRIOR APPLICATION NUMBER: 07/393,621
PRIOR FILING DATE: 1989-08-14
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-09-816-737-1
```

```
Query Match          100.0%; Score 80; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  GTPGPGIAGORGV 15
        |||
Db      1  GTPGPGIAGORGV 15
```

```
RESULT 3
US-09-935-417-1
Sequence 1, Application US/09935417
Patent No. US20020062145A1
GENERAL INFORMATION:
APPLICANT: Rudakov, Leon V.
APPLICANT: Imran, Mir A.
APPLICANT: Dinh, Linh
APPLICANT: Davidian, Ara
APPLICANT: Larkin, Kevin
TITLE OF INVENTION: Composite Expandable Device with Polymeric Covering and Bioactive
FILE REFERENCE: 52200-8006.US01
CURRENT APPLICATION NUMBER: US/09/935,417
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: US 09/385,691
PRIOR FILING DATE: 1999-08-30
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: portion of a1 chain of collagen
US-09-935-417-1
```

```
Query Match          100.0%; Score 80; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  GTPGPGIAGORGV 15
        |||
Db      1  GTPGPGIAGORGV 15
```

```
RESULT 4
US-10-133-289-1
Sequence 1, Application US/10133289
Publication No. US20030077825A1
GENERAL INFORMATION:
APPLICANT: Rajendra S. Bhatnagar
APPLICANT: Jing, Jing Qian
TITLE OF INVENTION: Structures Useful for Bone Engineering
FILE REFERENCE: UCAL224
CURRENT APPLICATION NUMBER: US/10/133,289
CURRENT FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: US/09/561,554
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Peptide
US-10-133-289-1
```

```
Query Match          100.0%; Score 80; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  GTPGPGIAGORGV 15
        |||
Db      1  GTPGPGIAGORGV 15
```

```
RESULT 5
US-10-176-401-1
Sequence 1, Application US/10176401
Publication No. US20030103960A1
GENERAL INFORMATION:
APPLICANT: Philippart, Pierre
APPLICANT: Brasseur, Michele
TITLE OF INVENTION: Sealant and bone generating product
FILE REFERENCE: 402119
CURRENT APPLICATION NUMBER: US/10/176,401
CURRENT FILING DATE: 2002-06-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(15)
OTHER INFORMATION: residue 766 to 780 from human collagen type 1
PUBLICATION INFORMATION:
AUTHORS: Bhatnagar RS, Qian JY, Gough CA
TITLE: The role in cell binding of a beta-bend within the triple helical region
TITLE: in collagen alpha 1 (I) chain: structural and biological evidence for
TITLE: conformational tautomerism on fiber surface
JOURNAL: Journal of Biomolecular Structure and Dynamics
VOLUME: 14
ISSUE: 5
PAGES: 547-560
DATE: 1997-04-01
DATABASE ENTRY DATE:
RELEVANT RESIDUES: (766)..(780)
US-10-176-401-1
```

```
Query Match          100.0%; Score 80; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	GTPGPGGIAGQRCV	15
Db	1	GTPGPGGIAGQRCV	15

RESULT 6
US-10-017-193-1

```

? Sequence 1, Application US/10017193
? Publication No. US20030113478A1
? GENERAL INFORMATION:
?
? APPLICANT: Dang, Mai Hung
? APPLICANT: Chu, Phillip
? TITLE OF INVENTION: Surface Coating Method and Coated Device
? FILE REFERENCE: 52200-8010
? CURRENT APPLICATION NUMBER: US/10/017,193
? CURRENT FILING DATE: 2001-12-12
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? LENGTH: 15
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
?
? OTHER INFORMATION: attachment peptide from collagen
? US-10-017-193-1

```

Query Match	100.0%	Score 80;	DB 15;	Length 15;
Best Local Similarity	100.0%	Pred. No. 3.8e-05;		
Matches	15;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY 1 GTPGPGGIAGQGRVV 15
 Db 1 GTPGPGGIAGQGRVV 15

RESULT 7
US-10-279-991-3

```

Sequence 3, Application US/10279991
Publication No. US20030087315A1
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: FERTALA, ANDRZEJ
TITLE OF INVENTION: INHIBITORS OF COLLAGEN ASSEMBLY
FILE REFERENCE: 053844-5001-01
CURRENT APPLICATION NUMBER: US/10/279, 991
PRIORITY FILING DATE: 2002-10-24
PRIORITY APPLICATION NUMBER: 09/517, 866
PRIORITY FILING DATE: 2000-03-03
PRIORITY APPLICATION NUMBER: 60/058, 353
PRIORITY FILING DATE: 1997-09-10
PRIORITY APPLICATION NUMBER: PCT/US98/18838
PRIORITY FILING DATE: 1998-09-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 25
TYPE: PR1
ORGANISM: Homo sapiens
US-10-279-991-3

```

Query Match	100.08;	Score 80;	DB 15;	length 25;
Best Local Similarity	100.08;	Pred. No. 6.3e-05;		
Matches 15; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	GTPGPQGIAGQRCV	15
Db	9	GTPGPQGIAGQRCV	23

RESULT 8
US-10-058-124-18
; Sequence 18, Application US/10058124
; Publication No. US20030119058A1

GENERAL INFORMATION:

TITLE OF INVENTION: A Method for Assaying Collagen Fragments in Body Fluids, A Test Kit and Means for Carrying Out the Method and Use of the Method to Diagnose the Presence of Disorders Associated with the Metabolism of

NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/570,573

ATTORNEY/AGENT INFORMATION:

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237

INFORMATION FOR SEQ ID NO: 18:

LENGTH: 1341 amino acids

ORIGINAL SOURCE:

IMMEDIATE SOURCE:

SEQUENCE DESCRIPTION:	SEQ ID NO:
8-13A-19	18

US-10-058-124-18

Query Match	100.0%;	Score 80;	DB 15;	Length 1341;
Best Local Similarity	100.0%;	Pred. No. 0.0031;		
Matches 15; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	GTPGPGIAGQGRV	15
Db	823	GTPGPGIAGQGRV	83

RESULT 9

; Sequence 28, Application US/10301822
; Publication No. US20030148410A1

APPLICANT: Millennium Pharmaceuticals, Inc

APPLICANT: Guillemette, Tracy L.

APPLICANT: Schlegel, Robert

APPLICANT: Thibodeau, Stephen N.
Burgart, Lawrence J.

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

;; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;; FILE REFERENCE: MPM01-029P2RNM
;; CURRENT APPLICATION NUMBER: US/10/301,822
;; CURRENT FILING DATE: 2002-11-21
;; PRIOR APPLICATION NUMBER: US 60/339,971
;; PRIOR FILING DATE: 2001-12-10
;; PRIOR APPLICATION NUMBER: US 60/361,978
;; PRIOR FILING DATE: 2002-03-05
;; PRIOR APPLICATION NUMBER: US 60/381,988
;; PRIOR FILING DATE: 2002-05-20
;; NUMBER OF SEQ ID NOS: 228
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 28
;; LENGTH: 1464
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-10-301-822-28

Query Match 100.0%; Score 80; DB 12; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCVV 15
|||||

DB 947 GTPGPGIAGGRCVV 961

RESULT 10
US-10-060-036-159
;; Sequence 159, Application US/10060036
;; Publication No. US20030073144A1
;; GENERAL INFORMATION:
;; APPLICANT: Benson, Darin R.
;; APPLICANT: Kalos, Michael J.
;; APPLICANT: Lodes, Michael J.
;; APPLICANT: Persing, David H.
;; APPLICANT: Hepler, William T.
;; APPLICANT: Jiang, Yugu
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
;; FILE REFERENCE: 210121.566
;; CURRENT APPLICATION NUMBER: US/10/060,036
;; CURRENT FILING DATE: 2002-01-30
;; NUMBER OF SEQ ID NOS: 4560
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 159
;; LENGTH: 1464
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-060-036-159

Query Match 100.0%; Score 80; DB 15; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCVV 15
|||||

DB 947 GTPGPGIAGGRCVV 961

RESULT 11
US-10-171-311-36
;; Sequence 36, Application US/10171311
;; Publication No. US20030087270A1
;; GENERAL INFORMATION:
;; APPLICANT: Schlegel, Robert
;; APPLICANT: Chen, Yan
;; APPLICANT: Zhao, Xumei
;; APPLICANT: Monahan, John
;; APPLICANT: Kamatkar, Shubhangi
;; APPLICANT: Glatz, Karen
;; APPLICANT: Gannavairapu, Manjula

;; APPLICANT: Hoersch, Sebastian
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;; TITLE OF INVENTION: OF CERVICAL CANCER
;; FILE REFERENCE: MPT-035
;; CURRENT APPLICATION NUMBER: US/10/171,311
;; CURRENT FILING DATE: 2002-06-12
;; PRIOR APPLICATION NUMBER: US 60/298,159
;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/298,155
;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/335,936
;; PRIOR FILING DATE: 2001-11-14
;; NUMBER OF SEQ ID NOS: 238
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 36
;; LENGTH: 1464
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-171-311-36

Query Match 100.0%; Score 80; DB 15; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCVV 15
|||||

DB 947 GTPGPGIAGGRCVV 961

RESULT 12
US-10-216-705-21
;; Sequence 21, Application US/10216705
;; Publication No. US20030096973A1
;; GENERAL INFORMATION:
;; APPLICANT: Meristem Therapeutics, S.A.
;; TITLE OF INVENTION: Recombinant Collagens and Derived Proteins Produced by Plants,
;; TITLE OF INVENTION: Obtaining Such and Their Uses
;; FILE REFERENCE: 1149-3 DIV
;; CURRENT APPLICATION NUMBER: US/10/216,705
;; CURRENT FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: US 09/331,347
;; PRIOR FILING DATE: 1999-08-17
;; NUMBER OF SEQ ID NOS: 22
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 21
;; LENGTH: 1464
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-216-705-21

Query Match 100.0%; Score 80; DB 15; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCVV 15
|||||

DB 947 GTPGPGIAGGRCVV 961

RESULT 13
US-10-149-352-2
;; Sequence 2, Application US/10149352
;; Publication No. US20030105050A1
;; GENERAL INFORMATION:
;; APPLICANT: Berl, Rajinder
;; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
;; FILE REFERENCE: 06275-254US1
;; CURRENT APPLICATION NUMBER: US/10/149,352
;; CURRENT FILING DATE: 2002-06-10
;; PRIOR APPLICATION NUMBER: PCT/GB00/04741
;; PRIOR FILING DATE: 2000-12-12
;; PRIOR APPLICATION NUMBER: GB 9929487.8

;; PRIOR FILING DATE: 1999-12-15
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn Ver. 4.0
;; SEQ ID NO 2
;; LENGTH: 1464
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-149-352-2

Query Match 100.0%; Score 80; DB 15; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCVV 15
DB 947 GTPGPGIAGGRCVV 961

RESULT 14
US-10-177-293-65

;; Sequence 65, Application US/10177293
;; Publication No. US20030124128A1

;; GENERAL INFORMATION:

;; APPLICANT: Lillie, James

;; APPLICANT: Glatz, Karen

;; APPLICANT: Zhao, Xumei

;; APPLICANT: Gannavarpu, Manjula

;; APPLICANT: Kamatkar, Shubhangl

;; APPLICANT: Mertens, Maureen

;; APPLICANT: Myer, Vic

;; APPLICANT: Wang, Youzhen

;; APPLICANT: Hoersch, Sebastian

;; APPLICANT: Monahan, John

;; APPLICANT: Meyers, Rachel E.

;; APPLICANT: Bast Jr., Robert C.

;; APPLICANT: Hortobagyl, Gabriel N.

;; APPLICANT: Pusztai, Lajos

;; APPLICANT: Metic, Funda

;; APPLICANT: Sahin, Aysegul

;; APPLICANT: Mills, Gordon B.

;; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
;; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER

;; FILE REFERENCE: MRI-038

;; CURRENT APPLICATION NUMBER: US/10/177,293

;; PRIOR FILING DATE: 2002-06-21

;; PRIOR APPLICATION NUMBER: US 60/299,887

;; PRIOR FILING DATE: 2001-06-21

;; PRIOR APPLICATION NUMBER: US 60/301,572

;; PRIOR FILING DATE: 2001-06-27

;; PRIOR APPLICATION NUMBER: US 60/306,501

;; PRIOR FILING DATE: 2001-07-18

;; PRIOR APPLICATION NUMBER: US 60/325,002

;; PRIOR FILING DATE: 2001-09-25

;; PRIOR APPLICATION NUMBER: US 60/362,585

;; PRIOR FILING DATE: 2002-03-05

;; PRIOR APPLICATION NUMBER: US 60/xxx,xxx

;; PRIOR FILING DATE: 2002-05-14

;; NUMBER OF SEQ ID NOS: 506

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 65

;; LENGTH: 1464

;; TYPE: PRT

;; ORGANISM: Homo sapiens

US-10-177-293-65

Query Match 100.0%; Score 80; DB 15; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0.0034;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCVV 15
DB 947 GTPGPGIAGGRCVV 961

RESULT 15
US-09-895-674-1

;; Sequence 1, Application US/09895674

;; Publication No. US20030021821A1

;; GENERAL INFORMATION:

;; APPLICANT: Fertala, Andrzej

;; APPLICANT: Ko, Frank

;; TITLE OF INVENTION: Collagen and Collagen-like Peptide Containing Polymeric

;; FILE REFERENCE: DRE-0032

;; CURRENT APPLICATION NUMBER: US/09/895,674

;; PRIOR FILING DATE: 2001-06-28

;; PRIOR APPLICATION NUMBER: PCT/US01/

;; PRIOR FILING DATE: 2001-06-25

;; PRIOR APPLICATION NUMBER: 60/ 214,034

;; NUMBER OF SEQ ID NOS: 1

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 1

;; LENGTH: 234

;; TYPE: PRT

;; ORGANISM: Homo sapiens

US-09-895-674-1

Query Match 88.8%; Score 71; DB 11; Length 234;
Best Local Similarity 80.0%; Pred. No. 0.011;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCVV 15
DB 67 GTPGPGIAGGRCVV 81

Search completed: August 29, 2003, 18:30:43
Job time : 26 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2003, 18:20:35 ; Search time 40 Seconds
(without alignments)
36.063 Million cell updates/sec

Title: US-09-935-417-1

Perfect score: 80

Sequence: 1 GTPGPGIAGRGVY 15

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	779	1 CGB01S	collagen alpha 1(I)
2	80	100.0	1042	1 CGCH1S	collagen alpha 1(I)
3	80	100.0	1453	2 S21626	collagen alpha 1(I)
4	80	100.0	1464	1 CGH01S	collagen alpha 1(I)
5	71	88.8	1418	2 T45467	collagen alpha 1(I)
6	71	88.8	1419	2 A41182	collagen alpha 1(I)
7	71	88.8	1487	1 CGH06C	collagen alpha 1(I)
8	71	88.8	1487	2 B41182	collagen alpha 1(I)
9	68	85.0	1486	1 B40333	collagen alpha 1(I)
10	62	77.5	1763	2 S16366	collagen alpha 2(I)
11	59	73.8	618	2 S32436	collagen alpha 2(I)
12	58	72.5	290	2 A32249	collagen alpha 2(I)
13	58	72.5	1496	1 CGH02V	collagen alpha 2(V)
14	57	71.2	1497	2 I49607	procollagen type V
15	57	71.2	310	2 I50696	collagen alpha 1(I)
16	57	71.2	636	2 S41067	collagen alpha 1(I)
17	57	71.2	886	2 I50694	collagen alpha 1(I)
18	57	71.2	1049	1 CGB07S	collagen alpha 1(I)
19	57	71.2	1414	1 S23809	collagen alpha 1(I)
20	57	71.2	1464	2 S59856	collagen alpha 1(I)
21	57	71.2	1466	1 CGH07L	collagen alpha 1(I)
22	56	70.0	677	2 S23296	collagen alpha 2(I)
23	56	70.0	920	2 A45748	collagen alpha 1(V)
24	56	70.0	1366	1 CGH02S	collagen alpha 1(I)
25	56	70.0	1669	1 CGH04B	collagen alpha 1(I)
26	55	68.8	1669	1 CGMS4B	collagen alpha 1(I)
27	55	68.8	1051	2 A53330	collagen alpha 2(I)
28	55	68.8	1051	2 A35763	collagen alpha 2(C)
29	55	68.8	1492	2 A40333	collagen alpha 1(I)

30	54	67.5	298	2 T32371	hypothetical prote
31	53	66.2	266	2 T22706	hypothetical prote
32	53	66.2	287	2 T15779	hypothetical prote
33	53	66.2	299	2 T22705	hypothetical prote
34	53	66.2	305	2 T30165	hypothetical prote
35	53	66.2	307	2 T19582	hypothetical prote
36	53	66.2	673	1 CGH06C	collagen alpha 1(I)
37	53	66.2	680	1 CGH01D	collagen alpha 1(X)
38	53	66.2	751	2 S64741	cuticle collagen
39	53	66.2	775	2 A61228	collagen alpha 2(I)
40	53	66.2	959	2 S32605	collagen alpha 3(V)
41	53	66.2	1142	2 JX0369	collagen alpha 1(X)
42	53	66.2	1603	2 S23810	collagen alpha 1(X)
43	53	66.2	1752	2 A45407	collagen alpha 3(I)
44	53	66.2	1758	2 T29350	hypothetical prote
45	53	66.2	1759	2 T29351	collagen alpha 2(I)

ALIGNMENTS

RESULT 1
CGB01S
collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence_revision 31-Dec-1993 #text_change 31-Mar-2000
C:Accession: A91193; A91229; A91387; A91211; A91201; A91200; A43048; A02853
R:Rautenberg, J.; Timpl, R.; Furthmayr, H.
Eur. J. Biochem. 27, 231-237, 1972
A:Title: Structural characterization of N-terminal antigenic determinants in calf and
A:Reference number: A91193; MUID:7225534; PMID:4115172
A:Accession: A91193
A:Molecule type: protein
A:Residues: 1-19 <RAU>
A:Experimental source: skin
A>Note: the epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is co
R:Rietzek, P.P.; Kuehn, K.
Eur. J. Biochem. 52, 77-82, 1975
A:Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-brom
A:Reference number: A91229; MUID:76022320; PMID:1164916
A:Accession: A91229
A:Molecule type: protein
A:Residues: 20-145 <RTD>
A:Experimental source: skin
A:Note: Lys-103 is hydroxylated and binds glucosylgalactose
R:Rietzek, P.P.; Wendt, P.; Kell, I.; Kuehn, K.
FBBS Lett. 26, 74-76, 1972
A:Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from c
A:Reference number: A91387; MUID:73049499; PMID:4673951
A:Accession: A91387
A:Molecule type: protein
A:Residues: 146-294 <RT2>
A:Experimental source: skin
R:Rietzek, P.P.; Rexrodt, F.W.; Hopper, K.E.; Kuehn, K.
Eur. J. Biochem. 38, 396-400, 1973
A:Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7
A:Reference number: A91211; MUID:74086118; PMID:4359390
A:Accession: A91211
A:Molecule type: protein
A:Residues: 295-562 <RT3>
A:Experimental source: skin
R:Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.
Eur. J. Biochem. 30, 169-183, 1972
A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 resid
A:Reference number: A91201; MUID:73042276; PMID:4343808
A:Accession: A91201
A:Molecule type: protein
A:Residues: 563-675 <WEN>
A:Experimental source: skin
R:Rietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.
Eur. J. Biochem. 30, 163-168, 1972
A:Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB
A:Reference number: A91200; MUID:73042275; PMID:4343807

A:Accession: A91200
 A:Molecule type: protein
 A:Residues: 676-758 <F14>
 A:Experimental source: skin
 A:Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in position 726.
 R:Rateberg, J.; Fietzek, F.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.
 FEBS Lett. 21, 75-79, 1972
 A:Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of
 A:Reference number: A43048
 A:Accession: A43048
 A:Molecule type: protein
 A:Residues: 759-779 <RA2>
 A:Experimental source: skin
 C:Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydroxylated.
 C:Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated.
 C:Comment: The order of the eight CNBr peptides in the alpha 1(I) chain of bovine skin is 9, 149, 268, and 217 residues.
 C:Comment: The complete chain contains 1052 residues.
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coll: extracellular matrix; glycoprotein; pyroglutamic acid; trimer;
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match

Best Local Similarity 100.0%; Score 80; DB 1; Length 779;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GTPGPGIAGRGVY 15
 |||||

Db 509 GTPGPGIAGRGVY 523

RESULT 2

CGCHS
 collagen alpha 1(I) chain - chicken (tentative sequence) (fragments)

C:Species: Gallus gallus (chicken)
 C:Date: 12-Aug-1991 #sequence revision 06-Jul-1982 #text-change 31-Mar-2000
 C:Accession: A90458; A90181; A02857
 R:Highberger, J.H.; Corbett, C.; Dixit, S.N.; Yu, W.; Seyer, J.M.; Kang, A.H.; Gross, J.
 Biochemistry 21, 2048-2055, 1982
 A:Title: Amino acid sequence of chick skin collagen alpha 1(I)-C88 and the complete prime
 A:Reference number: A90458; MUID:82231995; PMID:7093229
 A:Accession: A90458
 A:Molecule type: protein
 A:Residues: 1-1036 <HIG>
 A:Experimental source: skin
 A:Note: This is the latest in a series of papers from these workers elucidating the sequence.
 R:Exre, D.R.; Glimcher, M.J.
 Biochem. Biophys. Res. Commun. 48, 720-726, 1972
 A:Title: Evidence for a previously undetected sequence at the carboxyterminus of the alpha 1(I) chain.
 A:Reference number: A90181; MUID:72243016; PMID:5047697
 A:Accession: A90181
 A:Molecule type: protein
 A:Residues: 1037-1042 <EYR>
 A:Experimental source: skin
 A:Note: Residues 1037-1042 above correspond to the carboxyl end of the protein.
 C:Comment: Lysines at positions 103, 700, 934, and 946 above may be hydroxylated in some species.
 C:Comment: Most of the prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated.
 C:Comment: Pro-1002 is the only 3-hydroxyproline and the only hydroxylated proline in position 1002.
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coll: extracellular matrix; glycoprotein; pyroglutamic acid; trimer;
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match

Best Local Similarity 100.0%; Score 80; DB 1; Length 1042;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GTPGPGIAGRGVY 15
 |||||

Db 785 GTPGPGIAGRGVY 799

RESULT 3

521626

collagen alpha 1(I) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence revision 25-Apr-1997 #text-change 13-Aug-1999
 C:Accession: S57243; S16374; A23982; I49559; I45557; S39789; I48300; S21626
 R:Li, S.W.; Khillan, J.; Prockop, D.J.
 Matrix Biol. 14, 593-595, 1994
 A:Title: The complete cDNA coding sequence for the mouse pro-alpha-1(I) chain of type I collagen.
 A:Reference number: S57243
 A:Accession: S57243
 A:Molecule type: mRNA
 A:Residues: 1-1453 <LIS>
 A:Cross-references: EMBL:008020; NID:9470673; PIDN:AAA8912.1; PID:9470674
 R:Mesaaranta, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.
 Biochim. Biophys. Acta 1089, 241-243, 1991
 A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNA
 A:Reference number: S16176; MUID:91274355; PMID:2054384
 A:Accession: S16374
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1442-1453 <MET>
 A:Cross-references: EMBL:X57981; NID:950484; PIDN:CAA1046.1; PID:950485
 R:French, B.T.; Lee, W.H.; Maul, G.G.
 Gene 39, 311-312, 1985
 A:Title: Nucleotide sequence of a cDNA clone for mouse proalpha1(I) collagen protein.
 A:Reference number: A23982; MUID:86137403; PMID:3841523
 A:Accession: A23982
 A:Molecule type: mRNA
 A:Residues: 518-1128 <FRE>
 A:Cross-references: GB:M14423; NID:9192261; PIDN:AAA3733.1; PID:9192262
 R:Monson, J.M.; Friedman, J.; McCarthy, B.J.
 Mol. Cell. Biol. 2, 1362-1371, 1982
 A:Title: DNA sequence analysis of a mouse pro-alpha-1(I) procollagen gene: Evidence for
 A:Reference number: I49559; MUID:83141374; PMID:6298597
 A:Accession: I49559
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 735-1130 <RES>
 A:Cross-references: GB:M17491; NID:9192263; PIDN:AAA37334.1; PID:9192264
 R:Harbers, K.; Kuehn, M.; Delius, H.; Jaenisch, R.
 Proc. Natl. Acad. Sci. U.S.A. 81, 1504-1508, 1984
 A:Title: Insertion of retrovirus into the first intron of alpha1(I) collagen gene leads
 A:Reference number: I49557; MUID:86170331; PMID:6324198
 A:Accession: I49557
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-25 <REA>
 A:Cross-references: GB:K01688; NID:9192246; PIDN:AAA37330.1; PID:9553881
 R:Penton, S.P.; Lamande, S.R.; Hannagan, M.; Stacey, A.; Jaenisch, R.; Bateman, J.F.
 Biochim. Biophys. Acta 1216, 469-474, 1993
 A:Title: Genomic sequence of mouse COL1A1 encoding the collagen propeptides.
 A:Reference number: S39789; MUID:94092741; PMID:8268229
 A:Accession: S39789
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54876; NID:950486; PIDN:CAA38657.1; PID:950487
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMB

Query Match 100.0%; Score 80; DB 2; Length 1453;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTPGPGIAGRGV 15
 |||||
 Db 936 GTPGPGIAGRGV 950

RESULT 4
 CGHUIS
 collagen alpha 1(I) chain precursor - human
 N.Alternate names: procollagen alpha 1(I) chain
 C.Species: Homo sapiens (man)
 C.Date: 12-Aug-1981 #sequence_revision 04-Oct-1996 #text_change 31-Dec-2000
 C.Accession: I60114; S01143; A93335; I55254; A39943; I55237; A35233; S09400; B90567; S115269; A29439; I53466; A02852; I37247
 R.D'Alessio, M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.
 Gene 67, 105-115, 1988

A.Title: Complete nucleotide sequence of the region encompassing the first twenty-five
 A.Reference number: I60114; MUID:88329734; PMID:2843432
 A.Accession: I60114
 A.Status: translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-369, 'L', 371-589 <DAL>
 A.Cross-references: GB:M20789; NID:9179593; PIDN:AA59373.1; PID:9179594
 R.Tromp, G.; Kulvanant, H.; Stacey, A.; Shikata, H.; Baldwin, C.T.; Jaenisch, R.; Prock
 Blochem. J. 253, 919-922, 1988
 A.Title: Structure of a full-length cDNA clone for the prepro-alpha-1(I) chain of human
 A.Reference number: S01143; MUID:89025644; PMID:318743
 A.Accession: S01143
 A.Molecule type: mRNA
 A.Residues: 1-472 <TRK>
 A.Cross-references: EMBL:X07884; NID:930015; PIDN:CAA3073.1; PID:930016; GB:M36546; NID
 A.Note: submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
 R.Chu, M.L.; de Wet, W.; Bernard, M.; Ding, J.F.; Morabito, M.; Myers, J.; Williams, C.;
 Nature 310, 337-340, 1984
 A.Title: Human proalpha1(I) collagen gene structure reveals evolutionary conservation of
 A.Reference number: A93335; MUID:84270697; PMID:6462220
 A.Accession: A93335
 A.Molecule type: DNA
 A.Residues: 1-58, 'Q', 60-181 <CHD>
 A.Cross-references: EMBL:X00820; NID:935657; PIDN:CAA25394.1; PID:935658
 R.Rosouw, C.M.S.; Vergeer, W.P.; du Plooy, S.J.; Bernard, M.P.; Ramirez, F.; de Wet, W.
 J. Biol. Chem. 262, 15151-15157, 1987
 A.Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene en
 A.Reference number: I55254; MUID:88033098; PMID:2822714
 A.Accession: I55254
 A.Status: translation not shown; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-45 <ROS>
 A.Cross-references: GB:J02829; NID:9180387; PIDN:AA51993.1; PID:9180388
 R.Bornstein, P.; McKay, J.; Morishima, J.K.; Devareyalu, S.; Gellinas, R.E.
 Proc. Natl. Acad. Sci. U.S.A. 84, 8869-8873, 1987
 A.Title: Regulatory elements in the first intron contribute to transcriptional control c
 A.Reference number: A39943; MUID:88097389; PMID:3480516
 A.Accession: A39943
 A.Molecule type: DNA
 A.Residues: 1-34 <BOR>
 A.Cross-references: GB:J03559; NID:9180876; PIDN:AA52052.1; PID:9553238
 R.Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F.
 J. Biol. Chem. 260, 2315-2320, 1985
 A.Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter s
 A.Reference number: I55237; MUID:85130970; PMID:2857713
 A.Accession: I55237
 A.Status: translation not shown; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-34 <CH2>
 A.Cross-references: GB:M10627; NID:9180383; PIDN:AA51992.1; PID:9553226
 R.Wirtz, M.N.; Keene, D.R.; Horst, R.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Hollist
 J. Biol. Chem. 265, 6312-6317, 1990
 A.Title: In vivo and in vitro noncovalent association of excised alpha1(I) amino-termin

rome, type VII.
 A.Reference number: A35233; MUID:90202908; PMID:2318855
 A.Accession: A35233
 A.Molecule type: protein
 A.Residues: 33-52 <MIR>
 A.Note: this propeptide fragment remained non-covalently bound to a defective, unclon
 R.Well, D.; d'Alessio, M.; Ramirez, F.; de Wet, W.; Cole, W.G.; Chan, D.; Bateman, J.
 EMBO J. 8, 1705-1710, 1989
 A.Title: A base substitution in the exon of a collagen gene causes alternative splici
 A.Reference number: S09400; MUID:89356643; PMID:2767050
 A.Accession: S09400
 A.Molecule type: mRNA
 A.Residues: 156-183 <MEI>
 R.Click, E.M.; Bornstein, P.
 Biochemistry 9, 4699-4706, 1970
 A.Title: Isolation and characterization of the cyanogen bromide peptides from the alp
 A.Reference number: A90567; MUID:71038625; PMID:5529814
 A.Contents: CNBR0-1, CNBR2, CNBR4, CNBR5
 A.Accession: B90567
 A.Molecule type: protein
 A.Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233,
 A.Experimental source: skin
 A.Note: evidence for 170-allysine
 R.Baetge, B.; Notbohm, H.; Diebold, J.; Lehmann, H.; Bodo, M.; Deutzmann, R.; Mueller
 Eur. J. Biochem. 192, 153-159, 1990
 A>Title: A critical crosslink region in human-bone-derived collagen type I. Specific
 A.Reference number: S11372; MUID:90382436; PMID:2169412
 A.Accession: S11372
 A.Molecule type: protein
 A.Residues: 175-187, 274-287, 'P', 289 <BAE>
 A.Note: sequence of collagen alpha 1(S)(I) isolated from bone after pepsin digestion
 R.Beck, S.B.; Scholz, P.M.; Amenta, P.S.; Constantinou, C.D.; Levi-Minzl, S.A.; Gonzal
 J. Biol. Chem. 266, 21827-21832, 1991
 A>Title: The substitution of arginine for glycine 85 of the alpha 1(I) procollagen ch
 A.Reference number: I55342; MUID:92042092; PMID:1718984
 A.Accession: I55342
 A.Status: translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 258-268, 11347-1357 <DEA>
 A.Cross-references: GB:S67495; NID:9239007; PIDN:AA820350.1; PID:9239008
 A.Note: sequences from the 5' and 3' ends only are shown; mutant sequence 263-Arg rep
 R.Morgan, P.H.; Jacobs, H.G.; Segrest, J.P.; Cunningham, L.W.
 J. Biol. Chem. 245, 5042-5048, 1970
 A>Title: Comparative study of glycopeptides derived from selected vertebrate collagen
 A.Reference number: A92069; MUID:71001508; PMID:4319110
 A.Accession: A92069
 A.Molecule type: protein
 A.Residues: 263-268 <MOR>
 A.Experimental source: skin
 A.Note: attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose to 5-hydroxylysine
 R.Labhard, M.E.; Hollister, D.W.
 Matrix 10, 124-130, 1990
 A>Title: Segmental amplification of the entire helical and telopeptide regions of the
 A.Reference number: S15989; MUID:90326017; PMID:2374517
 A.Accession: S15989
 A.Molecule type: mRNA
 A.Residues: 281-302, 402-420, 823-843, 925-944, 1026-1045, 1143-1162 <LAB>
 R.Wirtz, M.N.; Rao, V.H.; Glanville, R.W.; Labhard, M.E.; Pretorius, P.J.; de Vries,
 Connect. Tissue Res. 29, 1-11, 1993
 A>Title: A cysteine for glycine substitution at position 175 in an alpha 1 (I) chain
 A.Reference number: I52905; MUID:93339042; PMID:8339541
 A.Accession: I52905
 A.Status: translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 342-352, 'C', 354-359 <MI2>
 A.Cross-references: GB:S64717; NID:9408195; PIDN:AA827677.1; PID:9408196
 A.Note: mutant sequence from patient with osteogenesis imperfecta
 R.Bernard, M.P.; Chu, M.L.; Myers, J.C.; Ramirez, F.; Elkenberry, E.F.; Prockop, D.J.
 Biochemistry 22, 5213-5223, 1983
 A>Title: Nucleotide sequences of complementary deoxyribonucleic acids for the proalph
 A.Reference number: A90476; MUID:84080385; PMID:6689127
 A.Accession: A90476

A: Molecule type: mRNA
 A: Residues: 425-1250, 'X', 1252-1328, 'S', 1330-1390, 'X', 1392-1464 <BER>
 A: Cross-references: GB:K01228; NID:G180391; PIDN:AAA51995.1; PID:G180392
 A: Note: sequence partially completed for missing nucleotides by A29439
 R: Chu, M.L.; Gargiulo, V.; Williams, C.J.; Ramirez, F.
 J. Biol. Chem. 260, 691-694, 1985
 A: Title: Multixon deletion in an osteogenesis imperfecta variant with increased type II
 A: Reference number: A22161; MUID:85104934; PMID:2981843
 A: Accession: A22161
 A: Molecule type: DNA
 A: Residues: 472-594, 'R', 596-607 <CH3>
 A: Cross-references: GB:K03178; GB:K03179; NID:G179612; NID:G179613; PIDN:AAA51847.1; PID
 A: Note: the authors translated the codon CGT for residue 595 as Pro
 R: Wallis, G.A.; Starman, B.J.; Zinn, A.B.; Byers, P.H.
 Am. J. Hum. Genet. 46, 1034-1040, 1990
 A: Title: Variable expression of osteogenesis imperfecta in a nuclear family is explained
 A: Reference number: A35336; MUID:90252792; PMID:2339700
 A: Accession: A35336
 A: Molecule type: mRNA
 A: Residues: 710-720, 'E', 722-737, 'E', 739-745 <MAL>
 A: Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu
 R: Forlino, A.; Zolietzi, F.; Valli, M.; Pignatelli, P.F.; Cetta, G.; Brunelli, P.C.; Motte
 Hum. Mol. Genet. 3, 2201-2206, 1994
 A: Title: Severe (type II) osteogenesis imperfecta due to glycine substitutions in the c
 A: Reference number: 154365; MUID:95187161; PMID:7881420
 A: Accession: 154365
 A: Status: translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 746-766, 'S', 768-781 <FOR>
 A: Cross-references: GB:I47667; NID:G1009093; PIDN:AA859576.1; PID:G1009094
 R: Chesser, S.D.; Wallis, G.A.; Byers, P.H.
 J. Biol. Chem. 268, 18218-18225, 1993
 A: Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of ty
 A: Reference number: A47426; MUID:93352646; PMID:8349697
 A: Accession: A47426
 A: Molecule type: mRNA
 A: Residues: 1179-1276, 'H', 1278-1336, 1339-1387, 'R', 1389-1464 <CHE>
 A: Cross-references: GB:564556; NID:G407589; PIDN:AA827856.1; PID:G407590
 A: Note: sequence extracted from NCBI backbone (NCBIN:136444, NCBI:136445)
 A: Note: does not represent an experimentally determined sequence but three different mut
 A: Accession: B47426
 A: Molecule type: mRNA
 A: Residues: 1179-1464 <CH4>
 A: Experimental source: normal dermal fibroblast culture
 A: Accession: C47426
 A: Molecule type: mRNA
 A: Residues: 1179-1276, 'H', 1278-1464 <CH5>
 A: Experimental source: fetal cell 86-237
 A: Accession: D47426
 A: Molecule type: mRNA
 A: Residues: 1179-1336, 1339-1464 <CH6>
 A: Experimental source: fetal cell 86-146
 A: Accession: E47426
 A: Molecule type: mRNA
 A: Residues: 1179-1387, 'R', 1389-1464 <CH7>
 A: Experimental source: fetal cell 88-251
 R: Choh, D.H.; Apore, S.; Eyre, D.R.; Starman, B.J.; Andreasen, P.; Charbonneau, H.; Nid
 J. Biol. Chem. 263, 14605-14607, 1988
 A: Title: Substitution of Cysteine for Glycine within the Carboxyl-terminal Telopeptide c
 A: Reference number: 155269; MUID:89008319; PMID:3170557
 A: Accession: 155269
 A: Status: translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1187-1194, 'C', 1196-1220 <COH>
 A: Cross-references: GB:M23213; NID:G340842; PIDN:AA859363.1; PID:G499622
 A: Note: mutant sequence from a patient with mild osteogenesis imperfecta
 R: Meckel, J.K.; Raassina, M.; Virta, A.; Vuorio, E.
 Nucleic Acids Res. 16, 349, 1988
 A: Title: Human pro-alpha-1(I) collagen: cDNA sequence for the C-propeptide domain.

Query Match 100.0%; Score 80; DB 1; Length 1464;
 Best Local Similarity 100.0%; Pred. No. 0.00014;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTPPGGIAGQGRGV 15
 Db 947 GTPPGGIAGQGRGV 961

RESULT 5

T45467

collagen alpha 1(II) chain precursor [imported] - horse

N: Alternate names: type II collagen

C: Species: Equus caballus (domestic horse)

C: Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 04-Mar-2000

R: Accession: T45467

R: Richardson, D.W.; Dodge, G.R.

submitted to the EMBL Data Library, June 1996

A: Description: Cloning of equine type II collagen and modulation of its expression in

A: Reference number: 222977

A: Accession: T45467

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: mRNA

A: Residues: 1-1418 <PIC>

A: Cross-references: EMBL:U62528; PIDN:AA805773.1

C: Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

Query Match

88.8%; Score 71; DB 2; Length 1418;

Best Local Similarity 80.0%; Pred. No. 0.0036;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTPPGGIAGQGRGV 15
 Db 900 GTPPGGIAGQGRGV 914

RESULT 6

A41182

collagen alpha 1(II) chain precursor - mouse

C: Species: Mus musculus (house mouse)

C: Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 13-Aug-1999

R: Accession: A41182; A44885

R: Metsger, M.; Roman, D.; de Crombrughe, B.; Vuorio, E.

J. Biol. Chem. 266, 16862-16869, 1991

A: Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, a

A: Reference number: A41182; MUID:91358489; PMID:1885613

A: Accession: A41182

A: Status: preliminary; not compared with conceptual translation

A: Molecule type: DNA

A: Residues: 1-1419 <MET>

A: Cross-references: GB:M65161

R: Cheah, K.S.; Lau, E.T.; Au, P.K.; Tam, P.P.

Development 111, 945-953, 1991

A: Title: Expression of the mouse alpha 1(II) collagen gene is not restricted to carti

A: Reference number: A44885; MUID:91347939; PMID:1879363

A: Accession: A44885

A: Molecule type: DNA

A: Residues: 1-28 <CHE>

A: Cross-references: GB:563190; NID:G234368; PIDN:AA819627.1; PID:G234369

A: Note: sequence extracted from NCBI backbone (NCBIN:63190, NCBI:63192)

C: Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

C: Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; tr

F; 1191-1419/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match

88.8%; Score 71; DB 2; Length 1419;

Best Local Similarity 80.0%; Pred. No. 0.0036;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTPPGGIAGQGRGV 15
 Db 901 GTPPGGIAGQGRGV 915

QY 1 GTPPGGIAGQGRGV 15
 Db 901 GTPPGGIAGQGRGV 915

RESULT 7

CGH06C

collagen alpha 1(II) chain precursor [validated] - human
 N.Alternate names: procollagen alpha 1(II) chain
 M.Contains: chondrocalcin, collagen alpha 1(II) chain precursor splice form 1; collagen
 C.Species: Homo sapiens (man)
 C.Date: 28-May-1986 #sequence, revision 01-Sep-1995 #text, change 08-Dec-2000
 C.Accession: A35513; S06715; S24270; A24828; S06496; A35428; A30147; A33116; S64674; S637250; I37251; I37252; I37253; I37254; I55338; I59535; I61910
 R.Ryan, M.C.; Sierski, M.; Sandell, L.J.
 Genomics 8, 41-48, 1990
 A.Title: The human type II procollagen gene: identification of an additional protein-cod
 A.Reference number: A38513; MUID:91184811; PMID:2081599
 A.Accession: A38513
 A.Molecule type: DNA
 A.Residues: 1-103 <RYA>
 A.Cross-references: GB:M20299; NID:9180883; PIDN:AAA73873.1; PID:9180884
 R.Su, M.W.; Lee, B.; Ramirez, F.; Machado, M.; Horton, W.
 Nucleic Acids Res. 17, 9473, 1989
 A.Title: Nucleotide sequence of the full length cDNA encoding for human type II procolla
 A.Reference number: S06715; MUID:90067946; PMID:2587267
 A.Accession: S06715
 A.Molecule type: mRNA
 A.Residues: 1-28, 'R', '99-1487 <SU2>
 A.Cross-references: EMBL:X16468; NID:929515; PIDN:CAA34488.1; PID:929516
 A.Note: alternative splice form 1
 R.Vikkula, M.; Metsaeranta, M.; Syvaenen, A.C.; Ala-Kokko, L.; Vuorio, E.; Peltonen, L.
 Blochem. J. 285, 287-294, 1992
 A.Title: Structural analysis of the regulatory elements of the type-II procollagen gene.
 A.Reference number: S24270; MUID:92344585; PMID:1637314
 A.Accession: S24270
 A.Status: translation not shown
 A.Molecule type: DNA
 A.Residues: 1-28 <VIK>
 A.Cross-references: EMBL:X58709; GB:S40537; NID:935569
 A.Note: This translation is not annotated in Genbank entry HSPROCOE1, release 111.0
 R.Nunez, A.M.; Kohno, K.; Martin, G.R.; Yamada, Y.
 Gene 44, 11-16, 1986
 A.Title: Promoter region of the human pro-alpha-1(II)-collagen gene.
 A.Reference number: A24828; MUID:87031574; PMID:3021582
 A.Accession: A24828
 A.Molecule type: DNA
 A.Residues: 1-8, 'T', '10-28 <NUN>
 A.Cross-references: GB:M25698; NID:9180872; PIDN:AAA52051.1; PID:9553237
 R.Baldwin, C.T.; Reginaldo, A.M.; Smith, C.; Jimenez, S.A.; Prockop, D.J.
 Blochem. J. 262, 521-528, 1989
 A.Title: Structure of cDNA clones coding for human type II procollagen. The alpha-1(II)
 A.Reference number: S06496; MUID:90026318; PMID:2803268
 A.Accession: S06496
 A.Molecule type: mRNA
 A.Residues: 7-28, 'R', '99-157', 'P', '159-440', 'G', '442-456', 'E', '458-640', 'A', '642-831', 'PA', '834', 'F',
 A.Cross-references: EMBL:X16711; NID:930040; PIDN:CAA34683.1; PID:930041
 A.Note: alternative splice form 1
 R.Ryan, M.C.; Sandell, L.J.
 J. Biol. Chem. 265, 10334-10339, 1990
 A.Title: Differential expression of a cysteine-rich domain in the amino-terminal propept
 A.Reference number: A35428; MUID:90285153; PMID:2355003
 A.Accession: A35428
 A.Status: not compared with conceptual translation
 A.Molecule type: mRNA
 A.Residues: 27-81, 'V', '83-103 <RYA2>
 A.Note: alternative splice form 2; splicing appears to be under developmental regulation
 R.Su, M.W.; Benson-Chanda, V.; Vlassing, H.; Ramirez, F.
 Genomics 4, 438-441, 1989
 A.Title: Organization of the exons coding for pro alpha-1(II) collagen N-propeptide cont
 A.Reference number: A30147; MUID:89233138; PMID:2714801
 A.Accession: A30147
 A.Molecule type: DNA
 A.Residues: 104-157, 'P', '159-236 <SUN>
 A.Cross-references: GB:M23660; GB:M25655; GB:M25656; GB:M25730; GB:M2168; GE
 R.Ala-Kokko, L.; Baldwin, C.T.; Moskowitz, R.W.; Prockop, D.J.
 Proc. Natl. Acad. Sci. U.S.A. 87, 6565-6568, 1990
 A>Title: Single base mutation in the type II procollagen gene (COL2A1) as a cause of pr
 A.Reference number: A94227; MUID:90370826; PMID:1975693
 A.Accession: A33116

A.Molecule type: DNA
 A.Residues: 171-172, 'C', '174-175 <ALA>
 A.Note: mutant sequence from a family with family with primary generalized osteoartru
 R.Diab, M.; Wu, J.J.; Eyre, D.R.
 Biochem. J. 314, 327-332, 1996
 A.Title: Collagen type IX from human cartilage: a structural protein of Intermolecula
 A.Reference number: S64673; MUID:96195147; PMID:8660302
 A.Accession: S64674
 A.Molecule type: protein
 A.Residues: 188-189, 'X', '191-195;1224-1230, 'X', '1233-1236 <DIA>
 R.Franc, S.; Marzin, E.; Boutillon, M.M.; Lafont, R.; Lechene de la Porte, P.; Herpag
 Eur. J. Biochem. 234, 125-131, 1995
 A>Title: Immunobistochemical and biochemical analyses of 20000-25000-year-old fossil
 A.Reference number: S63514; MUID:96096730; PMID:8529631
 A.Accession: S63514
 A.Molecule type: protein
 A.Residues: 243-261;575-590;756-763, 'X', '765-779 <PRA>
 R.Tiller, G.E.; Weis, M.A.; Polunbo, P.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Ey
 Am. J. Hum. Genet. 56, 388-395, 1995
 A>Title: An RNA-splicing mutation (G-51YS20) in the type II collagen gene (COL2A1) in
 A.Reference number: I38867; MUID:95150028; PMID:7847372
 A.Accession: I38867
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 440, 'G', '442-456', 'E', '458-480, 'P', '482-509 <TILL>
 A.Cross-references: EMBL:U15195; NID:9557053; PIDN:AB60370.1; PID:9557054
 R.Ramirez, F.
 Submitted to the EMBL Data Library, December 1988
 A.Reference number: S04892
 A.Accession: S04892
 A.Molecule type: mRNA
 A.Residues: 501-676, 'A', '678-783, 'A', '785-831, 'PA', '834', 'F', '836-1214 <RAM>
 A.Cross-references: EMBL:X13783; NID:930037; PIDN:CAA32030.1; PID:930050
 R.Vikkula, M.; Peltonen, L.
 FEBS Lett. 250, 171-174, 1989
 A>Title: Structural analyses of the polymorphic area in type II collagen gene.
 A.Reference number: S05000; MUID:89325561; PMID:2753125
 A.Accession: S05000
 A.Molecule type: DNA
 A.Residues: 630-640, 'A', '642-785 <VIK2>
 A.Cross-references: EMBL:X16158; NID:929951; PIDN:CAA34278.1; PID:91335018; PIDN:CAA3
 PIDN:CAA34283.1; PID:91335023; PIDN:CAA34284.1; PID:91335024
 R.Bogert, R.; Tiller, G.E.; Weis, M.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre
 J. Biol. Chem. 267, 22522-22526, 1992
 A>Title: An amino acid substitution (G1985->S19) in the collagen alpha 1(II) chain
 A.Reference number: A44309; MUID:93054548; PMID:1429602
 A.Accession: A44309
 A.Status: nucleic acid sequence not shown; not compared with conceptual translation
 A.Molecule type: DNA
 A.Residues: 752-831, 'PA', '834', 'F', '836-1005, 'K', '1007-1036, 'Q', '1038-1052, 'E', '1054-1068, '
 A.Cross-references: GB:I00977; NID:9180812; PIDN:AB23914.1; PID:9258774
 A.Note: sequence extracted from NCBI backbone (NCBI:11773); parts of this sequence
 A>Note: this translation is not annotated and this publication is not cited in Genban
 A>Note: mutant sequence associated with perinatal lethal hypochondrogenesis
 R.Tiller, G.E.; Rimoin, D.L.; Murray, L.M.; Cohn, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3869-3893, 1990
 A>Title: Tandem duplication within a type II collagen gene (COL2A1) exon in an indivi
 A.Reference number: S16502; MUID:90251662; PMID:2339128
 A.Accession: S16502
 A.Molecule type: DNA
 A.Residues: 1164-1184, 'GPSGKGANGIPGT', '1185-1199 <TIL2>
 A.Cross-references: EMBL:M71216; NID:9180808; PIDN:AAA52037.1; PID:9180809
 A.Note: mutant sequence from a patient with spondyloepiphyseal dysplasia
 R.Cheal, K.S.E.; Stoker, N.G.; Griffin, J.R.; Grosfeld, F.G.; Solomon, E.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2555-2559, 1985
 A>Title: Identification and characterization of the human type II collagen gene (COL2A
 A.Reference number: A02858; MUID:85190534; PMID:3857598
 A.Accession: A02858
 A.Molecule type: DNA
 A.Residues: 1032-1056, 'N', '1058-1068, 'T', '1070-1487 <CHE>
 A.Cross-references: GB:J00116; NID:9180395; PIDN:AAA51997.1; PID:9180396
 R.Elima, K.; Vuorio, T.; Vuorio, E.
 Nucleic Acids Res. 15, 9499-9504, 1987

A:Title: Determination of the single polyadenylation site of the human pro-alpha-1(II) c
 A:Reference number: A27280; MUID:88067771; PMID:2825137
 A:Accession: A27280
 A:Molecule type: DNA; mRNA
 A:Residues: 1175-1487 <ELI>
 A:Cross-references: EMBL:X06268; NID:930096; PIDN:CAA2604.1; PID:930097
 A:Experimental source: fetal epiphyseal cartilage
 A:van der Rest, M.; Rosenberg, L.C.; Olsen, B.R.; Poole, A.R.
 Biochem. J. 237, 923-925, 1986
 A:Title: Chondrocalcin is identical with the C-propeptide of type II procollagen.
 A:Reference number: A57033; MUID:87099927; PMID:3800925
 A:Accession: A57033
 A:Molecule type: protein
 A:Residues: 'XE',1244-1246,'N',1248,'X',1250-1265;1295-1305;1395-1408 <VAN>
 A:Note: chondrocalcin identified as released collagen 1(II) chain carboxyl-terminal pro
 R:Strom, C.M.; Upholt, W.B.
 Nucleic Acids Res. 12, 1025-1038, 1984
 A:Title: Isolation and characterization of genomic clones corresponding to the human tyf
 A:Reference number: A21733; MUID:84118798; PMID:6320112
 A:Accession: A21733
 A:Molecule type: DNA
 A:Residues: 1245-1295 <STR1>
 A:Cross-references: EMBL:X00339; EMBL:X00298; NID:9394699; PIDN:CAA25092.1; PID:94378975
 A:Accession: B21733
 A:Molecule type: DNA
 A:Residues: 894-909,'PE' <STR2>
 A:Cross-references: GB:K01785; NID:930035; PIDN:CAA25082.1; PID:91335032
 R:Nunez, A.M.; Francomano, C.; Young, M.F.; Martin, G.R.; Yamada, Y.
 Biochem. J. 24, 6343-6348, 1985
 A:Title: Isolation and partial characterization of genomic clones coding for a human pro
 gene.
 A:Reference number: A24561; MUID:86104139; PMID:3002437
 A:Accession: A24561
 A:Molecule type: DNA
 A:Residues: 1296-1358 <NUN2>
 A:Cross-references: GB:M12048; NID:9180017
 A:Note: This translation is not annotated in GenBank entry HM0CCT2A, release 111.0
 R:Sangiorgi, F.O.; Benson-Chanda, V.; de Wet, W.J.; Sobel, M.E.; Tsipouras, P.; Ramirez,
 Nucleic Acids Res. 13, 2207-2225, 1985
 A:Title: Isolation and partial characterization of the entire human pro alpha 1(II) coll
 A:Reference number: I37249; MUID:85215609; PMID:2987845
 A:Accession: S59491
 A:Molecule type: DNA
 A:Residues: 7-28,'R',99-114;541-578;786-802;1055-1056,'N',1058-1068,'T',1070-1109;1200-1
 A:Accession: I84453
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 7-28 <SAN2>
 A:Cross-references: GB:M23759; NID:9180845; EMBL:X03320; GB:M24938; NID:930104
 A:Note: the GenBank PID is based on an incorrect reading frame
 A:Accession: I37250
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 541-560 <SAN3>
 A:Cross-references: EMBL:X02378; GB:M23870; NID:930107; PIDN:CAA26227.1; PID:929621
 A:Accession: I37251
 Query Match 88.8%; Score 71; DB 1; Length 1487;
 Best Local Similarity 80.0%; Pred. No. 0.0038;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GTPGPGIAGGRCV 15
 | |||||:|||||:|
 Db 969 GPGPGIAGGRCV 983
 RESULT 8
 B41182
 collagen alpha 1(II) chain precursor (long splice form) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 16-Jul-1999
 C:Accession: B41182

R:Meisner, M.; Toman, D.; de Crombrughe, B.; Viorio, E.
 J. Biol. Chem. 266, 16862-16869, 1991
 A:Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, a
 A:Reference number: A41182; MUID:91358489; PMID:1885613
 A:Accession: B41182
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-1487 <MEP>
 A:Cross-references: GB:M65161
 C:Superfamily: collagen alpha 1(II) chain; fibrillar collagen carboxyl-terminal homolo
 C:Keywords: alternative splicing; coiled coll; extracellular matrix; glycoprotein; tr
 F:33-91/Domain: von Willibrand factor type C repeat homology <VWC>
 F:1259-1487/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
 Query Match 88.8%; Score 71; DB 2; Length 1487;
 Best Local Similarity 80.0%; Pred. No. 0.0038;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GTPGPGIAGGRCV 15
 | |||||:|||||:|
 Db 969 GPGPGIAGGRCV 983
 RESULT 9
 B40333
 collagen alpha 1(II) chain precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: B40333
 R:Su, M.W.; Suzuki, H.R.; Bleker, J.T.; Solursh, M.; Ramirez, F.
 J. Cell Biol. 115, 565-575, 1991
 A:Title: Expression of two nonallelic type II procollagen genes during Xenopus laevis
 A:Reference number: A40333; MUID:92011898; PMID:1918153
 A:Accession: B40333
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1486 <SUA>
 A:Cross-references: GB:M63595
 C:Superfamily: collagen alpha 1(II) chain; fibrillar collagen carboxyl-terminal homolo
 C:Keywords: coiled coll; extracellular matrix; glycoprotein; trimer; triple helix
 F:37-96/Domain: von Willibrand factor type C repeat homology <VWC>
 F:1258-1486/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
 Query Match 85.0%; Score 68; DB 1; Length 1486;
 Best Local Similarity 73.3%; Pred. No. 0.011;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GTPGPGIAGGRCV 15
 | |||||:|||||:|
 Db 971 GPGPGIAGGRCV 985
 RESULT 10
 S16366
 collagen alpha 2(IV) chain precursor - pig roundworm
 C:Species: Ascaris suum (pig roundworm)
 C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 13-Aug-1999
 C:Accession: S16366
 R:Petit, J.; Kingston, I.B.
 J. Biol. Chem. 266, 16149-16156, 1991
 A:Title: The complete primary structure of a nematode alpha-2(IV) collagen and the pa
 A:Reference number: S16366; MUID:91340768; PMID:1714907
 A:Accession: S16366
 A:Molecule type: mRNA
 A:Residues: 1-1763 <JBI>
 A:Cross-references: GB:M67507; NID:9159648; PIDN:AAA18014.1; PID:9159649
 C:Genetics:
 A:introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/3
 C:Superfamily: collagen alpha 1(IV) chain
 C:Keywords: alternative splicing; basement membrane; cell binding; coiled coll; disul
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-1763/Product: collagen alpha 2(IV) chain #status predicted <MAT>
 F:27-42/Domain: non-collagenous NHI #status predicted <NHI>

F:43-1529/Domain: collagenous #status predicted <COL>
 F:197-199/Region: cell attachment (R-G-D) motif
 F:1530-1763/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>
 F:1530-1638/Domain: repeat NC1 #status predicted <NC1>
 F:1639-1763/Domain: repeat NC1 #status predicted <NC1>
 F:31,34,39,41,536,539/Disulfide bonds: interchain #status predicted
 F:126/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:1593-1599,1702-1709/Disulfide bonds: #status predicted

Query Match 77.5%; Score 62; DB 2; Length 1763;
 Best Local Similarity 66.7%; Pred. No. 0.12;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GTPGPGIAGRGV 15
 Db 82 GPGPGIKGGRGII 96

RESULT 11

S32436
 collagen alpha 2(IX) chain - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-May-1998
 C:Accession: S32436; S34487; S64673
 R:Peraelae, M.; Haeninen, M.; Haestbacka, J.; Elima, K.; Vuorio, E.
 FEBS Lett. 319, 177-180, 1993
 A:Title: Molecular cloning of the human alpha-2(IX) collagen cDNA and assignment of the
 A:Reference number: S32436; MUID:93202262; PMID:8454052
 A:Accession: S32436
 A:Molecule type: mRNA
 A:Residues: 1-618 <PER1>
 A:Cross-references: EMBL:M95610; NID:91054872
 R:Peraelae, M.; Haeninen, M.; Haestbacka, J.; Vuorio, E.
 submitted to the EMBL Data Library, March 1993
 A:Description: Molecular cloning of the human alpha-2 (IX) collagen cDNA and assignment
 A:Reference number: S34487
 A:Accession: S34487
 A:Molecule type: mRNA
 A:Residues: 1-26, 'QT', 29, 'S', 31-32, 'LM', 35-561, 'L', 563-578, 'P', 580-618 <PER2>
 A:Cross-references: EMBL:M95610; NID:91054872
 R:Diab, M.; Wu, J.J.; Eyre, D.R.
 Biochem. J. 314, 327-332, 1996
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecular c
 A:Reference number: S64673; MUID:96195147; PMID:8660302
 A:Accession: S64673
 A:Molecule type: protein
 A:Residues: 123-133, 'P', 135-137 <DIA>
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit o
 ed and subsequently O-glycosylated.
 C:Genetics:
 A:Gene: GDB:COL9A2
 A:Cross-references: GDB:138310; OMIM:120260
 A:Map position: 1p33-1p32.2
 C:Complex: type IX collagen may be a heterotrimer of one alpha I(IX) chain, one alpha 2
 C:Function:
 A:Description: structural component of extracellular fibrous polymer associated with tyf
 C:Superfamily: unassigned collagens
 C:Keywords: chondroitin sulfate proteoglycan; coiled coil; extracellular matrix; glycop

F:1-114/Domain: collagenous COL3 (fragment) #status predicted <COL3>
 F:115-131/Domain: non-collagenous NC3 #status predicted <NC3>
 F:133-470/Domain: collagenous COL2 #status predicted <COL2>
 F:471-500/Domain: non-collagenous NC2 #status predicted <NC2>
 F:501-615/Domain: collagenous COL1 #status predicted <COL1>
 F:616-618/Domain: non-collagenous NC1 (fragment) #status predicted <NC1>
 F:120/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 73.8%; Score 59; DB 2; Length 618;
 Best Local Similarity 71.4%; Pred. No. 0.13;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GTPGPGIAGRGV 14
 Db 444 GPGPGIAGRGV 457

RESULT 12

A32249
 collagen - sea urchin (Paracentrotus lividus) (fragment)
 C:Species: Paracentrotus lividus (common urchin)
 C>Date: 17-Aug-1989 #sequence_revision 17-Aug-1989 #text_change 19-Jan-1996
 C:Accession: A32249
 R:Saltia, B.; Butlice, G.; Gambino, R.
 Biochem. Biophys. Res. Commun. 158, 633-639, 1989
 A:Title: Isolation of a putative collagen-like gene from the sea urchin Paracentrotus
 A:Reference number: A32249; MUID:89149773; PMID:2537631
 A:Accession: A32249
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-290 <SAI>
 C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homolo
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix

Query Match 72.5%; Score 58; DB 2; Length 290;
 Best Local Similarity 70.9%; Pred. No. 0.09;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GTPGPGIAGRG 13
 Db 268 GPGPGIAGRG 280

RESULT 13

CGH2V
 collagen alpha 2(V) chain precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 31-Jul-1988 #sequence_revision 28-Jul-1995 #text_change 31-Dec-2000
 C:Accession: A31427; A54555; S43643; A25874; I55239; I59025; A25374; A30017
 R:Woodbury, D.; Benson-Chanda, V.; Ramirez, F.
 J. Biol. Chem. 264, 2735-2738, 1989
 A:Title: Amino-terminal propeptide of human pro-alpha2(V) collagen conforms to the st
 A:Reference number: A31427; MUID:89123368; PMID:2914927
 A:Accession: A31427
 A:Molecule type: mRNA
 A:Residues: 1-463 <MOO>
 A:Cross-references: GB:J04478; NID:9179697; PIDN:AAA51859.1; PID:9179698
 A:Experimental source: placenta
 R:Greenspan, D.S.; Lee, S.T.; Lee, B.S.; Hoffman, G.G.
 Gene Expr. 1, 29-39, 1991
 A:Title: Homology between alpha2(V) and alpha1(III) collagen promoters and evidence f
 A:Reference number: A54555; MUID:92314691; PMID:1820205
 A:Accession: A54555
 A:Molecule type: DNA
 A:Residues: 1-32 <GRE>

A:Cross-references: GB:M58529; NID:9180834; PIDN:AA41699.1; PID:9553235
 R:Moradl-Ameli, M.; Rousseau, J.C.; Klemen, J.P.; Champilaud, M.F.; Bouillon, M.M.;
 Eur. J. Biochem. 221, 987-995, 1994
 A:Title: Diversity in the processing events at the N-terminus of type-V collagen.
 A:Reference number: SA3642; MUID:94237164; PMID:8181482
 A:Accession: SA3643
 A:Molecule type: protein
 A:Residues: 288-291, 'P', 293-294, 'X', 296-297, 606, 'X', 608-617 <MOR>
 R:Well, D.; Bernard, M.; Garcano, S.; Ramirez, F.
 Nucleic Acids Res. 15, 181-198, 1987
 A:Title: The pro alpha 2(V) collagen gene is evolutionarily related to the major fibr
 A:Reference number: A25874; MUID:87146331; PMID:3025669
 A:Accession: A25874
 A:Molecule type: mRNA; DNA
 A:Residues: 398-1496 <WEI>
 A:Cross-references: GB:X04758; NID:929588; PIDN:CAA28454.1; PID:91340175
 A:Experimental source: rhabdomyosarcoma cell line
 R:Myers, J.C.; Loidl, H.R.; Stollé, C.A.; Seyer, J.M.
 J. Biol. Chem. 260, 5533-5541, 1985
 A:Title: Partial covalent structure of the human alpha 2 type V collagen chain.
 A:Reference number: I55239; MUID:85182703; PMID:2985598
 A:Accession: I55239
 A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1002-1226 <RES>
A:Cross-references: GB:M10956; NID:g180427; PIDN:AAA52007.1; PID:g180428
A:Note: part of this sequence were determined by protein sequencing
R:Emmanuel, B.S.; Cammuzzaro, L.A.; Seyer, J.M.; Myers, J.C.
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985
A:Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm
A:Reference number: 159025; MUID:85216505; PMID:3858826
A:Accession: 159025
A:Status: translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1003-1034 <RES>
A:Cross-references: GB:M11135; NID:g179693; PIDN:AAA51857.1; PID:g179694
A:Note: part of this sequence were determined by protein sequencing
R:Myers, J.C.; Ioldi, H.R.; Seyer, J.M.; Dion, A.S.
J. Biol. Chem. 260, 11216-11222, 1985
A:Title: Complete primary structure of the human alpha-2 type V procollagen COOH-terminal
A:Reference number: A25374; MUID:85289337; PMID:2411731
A:Accession: A25374
A:Molecule type: mRNA
A:Residues: 1227-1417, 'T', 1419-1437, 'S', 1439-1496 <MYE>
A:Cross-references: GB:M11718; NID:g180912; PIDN:AAA52058.1; PID:g180913
A:Experimental source: normal fibroblasts
R:Jaispours, P.; Schwartz, R.C.; Liddell, A.C.; Salkeid, C.S.; Weil, D.; Ramirez, F.
Genomics 3, 275-277, 1988
A:Title: Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located on
A:Reference number: A30017; MUID:89138450; PMID:3224983
A:Accession: A30017
A:Molecule type: DNA
A:Residues: 1449-1463, 'E', 1465-1495, 'A' <TST>
A:Cross-references: GB:U03051; NID:g179695; PIDN:AAA51858.1; PID:g179696
A:Note: the authors translated the codon GAA for residue 1460 as Gln, and GAG for residue
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
are 5-hydroxylated and subsequently O-glycosylated.
C:Comment: The amino-terminal propeptide domain appears not to be completely cleaved.
C:Genetics:
A:Gene: GDB:COL5A2
A:Cross-references: GDB:119064; OMIM:120190
A:Map position: 2q31-2q31
A:Intons: 33/1; 812/3; 830/3; 848/3; 884/3; 902/3; 922/3; 974/3; 1046/3; 1064/3; 1448/3
C:Complex: type V collagen may be a homotrimer of alpha 1(V) chains (see PIR:GCH1V), a
alpha 2(V) chain and one alpha 3(V) chain, initially linked by disulfide bonds among the
length, is formed with desmosine cross-links made from lysine and allysine residues
C:Function:
A:Description: structural component of extracellular fibrous polymer associated with cell
A:Note: may play a role in controlling the lateral growth of collagen I fibrils
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-150/Product: collagen alpha 2(V) chain #status predicted <MAT>
F:127-150/Domain: amino-terminal propeptide (uncleaved) #status predicted <NMP>
F:27-108/Region: nonhelical
F:40-99/Domain: von Willebrand factor type C repeat homology <WCV>
F:109-186/Region: helical
F:187-208/Region: nonhelical
F:209-1225/Region: helical
F:503-505/Region: cell attachment (R-G-D) motif
F:941-943/Region: cell attachment (R-G-D) motif
F:1064-1066/Region: cell attachment (R-G-D) motif
F:1067-1068/Region: cell attachment (R-G-D) motif
F:1097-1099/Region: cell attachment (R-G-D) motif
F:1124-1126/Region: cell attachment (R-G-D) motif
F:1133-1135/Region: cell attachment (R-G-D) motif
F:1225-1250/Region: carboxyl-terminal nonhelical telopeptide
F:1251-1496/Domain: carboxyl-terminal propeptide #status predicted <CPP>
F:1269-1496/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
F:2/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:193-194/Cleavage site: Ala-Gln (procollagen N-endopeptidase) #status predicted
F:194/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:201/Modified site: allysine (lys) #status predicted
F:290-293-296-608-614-1004-1007-1013-1028-1034/Modified site: 4-hydroxyproline (Pro) #st
F:299-1139/Modified site: 5-hydroxylysine (lys) #status predicted
F:299-1139/Binding site: carbohydrate (lys) (covalent) #status predicted

F:1025/Modified site: 5-hydroxylysine (lys) #status experimental
F:1250-1251/Cleavage site: Glu-Asp (procollagen C-endopeptidase) #status predicted
F:1259-1397/Binding site: carbohydrate (asn) (covalent) #status predicted
F:1293-1299-1325/Disulfide bonds: interchain #status predicted
F:1333-1494-1402-1447/Disulfide bonds: #status predicted

Query Match 72.5%; Score 58; DB 1; Length 1496;
Best Local Similarity 66.7%; Pred. No. 0.45;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 GTPPGGAGGAGV 15
||||| |||||
Db 978 GPPGAGTTCGRCIV 992

RESULT 14
149607

procollagen type V alpha 2 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999

C:Accession: 149607

R:Andrikopoulos, K.; Suzuki, H.R.; Solursh, M.; Ramirez, F.
Dev. Dyn. 195, 113-120, 1992

A:Title: Localization of pro-alpha 2(V) collagen transcripts in the tissues of the de

A:Reference number: 149607; MUID:93214071; PMID:1297453

A:Accession: 149607

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: mRNA

A:Residues: 1-1497 <RES>

A:Cross-references: GB:I02918; NID:g309180; PIDN:AAA37440.1; PID:g309181

C:Genetics:

A:Gene: COL5A-2

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

F:39-98/Domain: von Willebrand factor type C repeat homology <WCV>

F:120-1497/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match

Best Local Similarity 72.5%; Score 58; DB 2; Length 1497;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 GTPPGGAGGAGV 15
||||| |||||
Db 979 GPPGAGTTCGRCIV 993

RESULT 15
150696

collagen alpha 1(III) chain - chicken (fragment)

C:Species: Gallus gallus (chicken)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Aug-1999

C:Accession: 150696

R:Neh, H.D.; Niu, Z.; Adams, S.L.
J. Biol. Chem. 269, 16443-16448, 1994

A:Title: An alternative transcript of the chick type III collagen gene that does not

A:Reference number: A54041; MUID:94266842; PMID:8206952

A:Accession: 150696

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: mRNA

A:Residues: 1-310 <RNA>

A:Cross-references: EMBL:U07974; NID:g520456; PIDN:AAA83409.1; PID:g537433

C:Genetics:

A:Gene: COL3A1

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

Query Match

Best Local Similarity 71.2%; Score 57; DB 2; Length 310;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GTPPGGAGGAGV 13
||||| |||||
Db 275 GPPGAGGAGG 287

Search completed: August 29, 2003, 18:27:51
Job time : 40 secs

THIS PAGE BLANK (us970)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2003, 17:51:44 ; Search time 23 Seconds
(without alignments)
30.670 Million cell updates/sec

Title: US-09-935-417-1
Perfect score: 80
Sequence: 1 GTPGPGGIAGRGVY 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	779	1 CA11_BOVIN	P02453 bos taurus
2	80	100.0	1453	1 CA11_CHICK	P02457 gallus galli
3	80	100.0	1453	1 CA11_MOUSE	P11087 mus musculus
4	80	100.0	1460	1 CA11_CANFA	O9x517 canis famli
5	80	100.0	1464	1 CA11_HUMAN	P02452 homo sapien
6	71	88.8	1418	1 CA12_HUMAN	P02458 homo sapien
7	71	88.8	1459	1 CA12_MOUSE	P28481 mus musculu
8	62	77.5	1763	1 CA24_ASCSU	P27393 ascaris suu
9	59	73.8	689	1 CA29_HUMAN	O14055 homo sapien
10	58	72.5	1362	1 CA21_CHICK	P02467 gallus galli
11	58	72.5	1496	1 CA25_HUMAN	P05997 homo sapien
12	57	71.2	636	1 CA13_RAT	P13941 rattus norv
13	57	71.2	1049	1 CA13_BOVIN	P04258 bos taurus
14	57	71.2	1262	1 CA13_CHICK	P12105 gallus galli
15	57	71.2	1464	1 CA13_MOUSE	P08121 mus musculu
16	57	71.2	1466	1 CA13_HUMAN	P02461 mus sapien
17	56	70.0	353	1 CA29_CHICK	P12108 gallus galli
18	56	70.0	526	1 CA21_RABIT	O28668 oryctolagus
19	56	70.0	1366	1 CA21_CANFA	O46392 canis famli
20	56	70.0	1366	1 CA21_HUMAN	P08123 homo sapien
21	56	70.0	1669	1 CA14_HUMAN	P02462 homo sapien
22	56	70.0	1669	1 CA14_MOUSE	P02463 mus musculu
23	54	67.5	747	1 CA12_MOUSE	P02459 bos taurus
24	53	66.2	266	1 YXWK_CAEEL	O21184 caenorhabdi
25	53	66.2	680	1 CA1A_HUMAN	O00392 homo sapien
26	53	66.2	1143	1 CA1F_HUMAN	O14993 homo sapien
27	53	66.2	1603	1 CA1F_HUMAN	O07092 homo sapien
28	53	66.2	1758	1 CA24_CAEEL	P11140 caenorhabdi
29	52	65.0	122	1 CA12_RAT	O05539 rattus norv
30	52	65.0	321	1 CA13_BOVIN	P42916 bos taurus
31	52	65.0	369	1 P5PD_BOVIN	P33246 bos taurus
32	52	65.0	371	1 C146_BOVIN	O8mh29 bos taurus
33	52	65.0	520	1 MRCO_HUMAN	O9new3 homo sapien

34	52	65.0	921	1 CA19_MOUSE	Q05722 mus musculu
35	52	65.0	1364	1 CA21_BOVIN	P02465 bos taurus
36	52	65.0	1736	1 CA2B_HUMAN	P13942 homo sapien
37	52	65.0	1736	1 CA2B_MOUSE	O64739 mus musculu
38	52	65.0	1804	1 CA1B_MOUSE	O61245 mus musculu
39	52	65.0	1806	1 CA1B_HUMAN	P12107 homo sapien
40	51	63.7	671	1 CA11_RAT	P02454 rattus norv
41	51	63.7	911	1 CA1B_BOVIN	O28083 bos taurus
42	51	63.7	1356	1 CA21_ONCMY	O93484 oncorhynch
43	51	63.7	1707	1 CA24_MOUSE	P08122 mus musculu
44	51	63.7	1712	1 CA24_HUMAN	P08572 homo sapien
45	51	63.7	1775	1 CA14_DROME	P08120 drosophila

ALIGNMENTS

RESULT 1	ID	CA11_BOVIN	STANDARD:	PRT:	779 AA.
AC	P02453;	21-JUL-1986 (Rel. 01, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Collagen alpha 1(I) chain (Fragments).				
CN	COL1A1.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OX	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxId=9913;				
RN	[1]				
RP	SEQUENCE OF 1-19.				
RA	MEDLINE=72255334; PubMed=4115172;				
RX	Rautenberg J., Jimpl R., Furtmayr H.;				
RT	"Structural characterization of N-terminal antigenic determinants in				
RT	calf and human collagen."				
RL	Eur. J. Biochem. 27:231-237(1972).				
RL	[2]				
RP	SEQUENCE OF 20-145.				
RX	MEDLINE=76022320; PubMed=1164916;				
RA	Fietzek P.P., Kuehn K.;				
RT	"The covalent structure of collagen: amino-acid sequence of the				
RT	cyano-gen-bromide peptides alpha-1-CB2, alpha-1-CB4 and alpha-1-CB5				
RT	from calf-skin collagen."				
RL	Eur. J. Biochem. 52:77-82(1975).				
RL	[3]				
RP	SEQUENCE OF 146-294.				
RX	MEDLINE=73049499; PubMed=4673951;				
RA	Fietzek P.P., Wendt P., Kell I., Kuehn K.;				
RT	"The covalent structure of collagen: amino acid sequence of alpha-1-				
RT	CB3 from calf skin collagen."				
RL	FEBS Lett. 26:74-76(1972).				
RL	[4]				
RP	SEQUENCE OF 295-562.				
RX	MEDLINE=74086118; PubMed=4359390;				
RA	Fietzek P.P., Rexrodt F.W., Hopper K.E., Kuehn K.;				
RT	"The covalent structure of collagen. 2. The amino-acid sequence of				
RT	alpha-1-CB7 from calf-skin collagen."				
RL	Eur. J. Biochem. 38:396-400(1973).				
RL	[5]				
RP	SEQUENCE OF 563-675.				
RX	MEDLINE=73042276; PubMed=4343808;				
RA	Wendt P., Mark K.V.D., Rexrodt F., Kuehn K.;				
RT	"The covalent structure of collagen. The amino-acid sequence of the				
RT	112-residues. Amino-terminal part of peptide alpha-1-CB6 from calf-				
RT	skin collagen."				
RL	Eur. J. Biochem. 30:169-183(1972).				
RL	[6]				
RP	SEQUENCE OF 676-751.				
RX	MEDLINE=73042275; PubMed=4343807;				
RA	Fietzek P.P., Rexrodt F.W., Wendt P., Stark M., Kuehn K.;				
RT	"The covalent structure of collagen. Amino-acid sequence of peptide				

DR PROSITE: PS50184; WVFC_2; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 22
 FT PROPEP 23 151 AMINO-TERMINAL PROPEPTIDE.
 FT CHAIN 152 1205 COLLAGEN ALPHA 1(I) CHAIN.
 FT PROPEP 1206 1453 C-TERMINAL PROPEPTIDE.
 FT DOMAIN 31 89 WVFC.
 FT MOD_RES 152 152 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 254 254 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 851 851 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 1081 1081 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 1097 1097 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 1153 1153 HYDROXYLATION.
 FT CONFLICT 1187 1187 F->L (IN REF. 5).
 FT CONFLICT 1441 1441 O->H (IN REF. 6).
 SQ SEQUENCE 1453 AA; 137789 MM; 3BC6152134271F4D CRC64;
 Query Match 100.0%; Score 80; DB 1; Length 1453;
 Best Local Similarity 100.0%; Pred. No. 0.00022;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPGPGIAGGRCVV 15
 |||||
 DB 936 GTPGPGIAGGRCVV 950

RESULT 3
 CALL_MOUSE STANDARD; PRT; 1453 AA.
 ID CALL_MOUSE
 AC P11087; O60635;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE Collagen alpha 1(I) chain precursor.
 GN COL1A1 OR COLA1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N;
 RX MEDLINE=96033240; PubMed=8535610;
 RA Li S.W., Khillan J., Prockop D.J.;
 RT "The complete cDNA coding sequence for the mouse pro alpha 1(I) chain
 of type I procollagen.";
 RL Matrix Biol. 14:593-595(1995).
 RN [2]
 RP SEQUENCE OF 518-1128 FROM N.A.
 RX MEDLINE=86137403; PubMed=3841523;
 RA French B.T., Lee W.-H., Maul G.G.;
 RT "Nucleotide sequence of a cDNA clone for mouse pro alpha 1(I)
 collagen protein.";
 RL Gene 39:311-312(1985).
 RN [3]
 RP SEQUENCE OF 735-1130 FROM N.A.
 RX MEDLINE=83141374; PubMed=6298597;
 RA Monson J.M., Friedman J., McCarthy B.J.;
 RT "DNA sequence analysis of a mouse pro alpha 1(I) procollagen gene:
 evidence for a mouse B1 element within the gene.";
 RL Mol. Cell. Biol. 2:1362-1371(1982).
 RN [4]
 RP SEQUENCE OF 735-878 AND 1005-1058 FROM N.A.
 RX MEDLINE=83157109; PubMed=6219867;
 RA Monson J.M., McCarthy B.J.;
 RT "Identification of a Balb/c mouse pro alpha 1(I) procollagen gene:
 evidence for insertions or deletions in gene coding sequences.";
 RL DNA 1:59-69(1981).
 RN [5]
 RP SEQUENCE OF 1442-1453 FROM N.A.
 RX MEDLINE=88124276; PubMed=3340560;
 RA Mooslehner K., Harbers K.;

FT "Two mRNAs of mouse pro alpha 1(I) collagen gene differ in the size
 RT of the 3'-untranslated region.";
 RL Nucleic Acids Res. 16:773-773(1988).
 CC -I- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
 CC (FIBRILLAR FORMING COLLAGEN).
 CC -I- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
 CC -I- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
 CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
 CC HYDROXYAPATITE.
 CC -I- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -I- SIMILARITY: Contains 1 WVFC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

CC EMBL; U08020; AAA88912.1; -;
 DR EMBL; X15896; CAA33904.1; -;
 DR EMBL; M14423; AAA37333.1; -;
 DR EMBL; M17491; AAA37334.1; -;
 DR EMBL; X06753; CAA29927.1; -;
 DR EMBL; X03036; AAA37332.1; -;
 DR EMBL; K03029; AAA37332.1; JOINED.
 DR EMBL; K03030; AAA37332.1; JOINED.
 DR EMBL; K03031; AAA37332.1; JOINED.
 DR EMBL; K03032; AAA37332.1; JOINED.
 DR EMBL; K03033; AAA37332.1; JOINED.
 DR EMBL; K03034; AAA37332.1; JOINED.
 DR EMBL; K03035; AAA37332.1; JOINED.
 DR PIR; S57243; S21626.
 DR MGD; MGI:88467; Colla1.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR000885; Fib_collagen_C.
 DR InterPro; IPR001007; WVFC_C.
 DR Pfam; PF01410; COLFI; 1.
 DR Pfam; PF01391; Collagen; 18.
 DR ProDom; PD000007; C1g_helix; 1.
 DR ProDom; PD002078; Fib_collagen_C; 1.
 DR SMART; SM00038; COLFI; 1.
 DR SMART; SM00214; WVFC; 1.
 DR PROSITE; PS01208; WVFC_1; 1.
 DR PROSITE; PS50184; WVFC_2; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen; Signal.
 FT SIGNAL 1 22
 FT PROPEP 23 151 AMINO-TERMINAL PROPEPTIDE.
 FT CHAIN 152 1207 COLLAGEN ALPHA 1(I) CHAIN.
 FT PROPEP 1208 1453 CARBOXYL-TERMINAL PROPEPTIDE.
 FT DOMAIN 29 87 WVFC.
 FT DOMAIN 152 167 NONHELICAL REGION (N-TERMINAL).
 FT DOMAIN 168 181 TRIPLE-HELICAL REGION.
 FT DOMAIN 1182 1207 NONHELICAL REGION (C-TERMINAL).
 FT CARBOHD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 1354 1354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 734 736 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1082 1084 CELL ATTACHMENT SITE (POTENTIAL).
 FT CONFLICT 1450 1450 A->V (IN REF. 5).
 SQ SEQUENCE 1453 AA; 137944 MM; 3B802E535DF81608 CRC64;

Query Match 100.0%; Score 80; DB 1; Length 1453;
 Best Local Similarity 100.0%; Pred. No. 0.00022;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTPGPGIAGGRCVV 15
 |||||
 DB 936 GTPGPGIAGGRCVV 950

```

RESULT 4
CALL_CANFA STANDARD: PRT: 1460 AA.
ID CALL_CANFA
AC 09XSJ7:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Collagen alpha 1(I) chain precursor.
GN COL1A1.
OS Canis familiaris (Dog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Campbell B.G., Wootton J.A.M., McLeod J.N., Minor R.R.;
RL "Sequence of normal canine COL1A1 cDNA.";
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC (FIBRILLAR FORMING COLLAGEN).
CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
CC -1- PM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- SIMILARITY: Contains 1 WMFC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF153062; AAD34619.1;
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000885; Fib_collagen_C.
DR InterPro: IPR001007; WMFC_C.
DR Pfam: PF01410; COLFI. 1.
DR Pfam: PF01391; Collagen; 18.
DR ProDom: PD000007; Ctg_helix; 2.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR SMART: SM00214; WMFC; 1.
DR PROSITE: PS01208; WMFC_1; 1.
DR PROSITE: PS50184; WMFC_2; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Collagen; Signal.
FT SIGNAL 1 22
FT PROPEP 23 157 BY SIMILARITY.
FT CHAIN 158 1214 AMINO-TERMINAL PROPEPTIDE.
FT PROPEP 1215 1460 COLLAGEN ALPHA 1(I) CHAIN.
FT DOMAIN 34 92 CARBOXYL-TERMINAL PROPEPTIDE.
FT DOMAIN 158 174 WMFC.
FT DOMAIN 175 1188 NONHELICAL REGION (N-TERMINAL).
FT DOMAIN 1189 1214 TRIPLE-HELICAL REGION.
FT SITE 741 743 NONHELICAL REGION (C-TERMINAL).
FT SITE 1089 1091 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 1361 1361 CELL ATTACHMENT SITE (POTENTIAL).
FT SEQUENCE 1460 AA: 138762 MW: 585674D2570697 CRC64;
SQ
Query Match 100.0%; Score 80; DB 1; Length 1460;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTPGPGIAGRGVY 15
 DB 943 GTPGPGIAGRGVY 957

RESULT 5
 CALL_HUMAN STANDARD: PRT: 1464 AA.

```

AC P02452; Q14037; Q15176;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Collagen alpha 1(I) chain precursor.
GN COL1A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-472 FROM N.A.
RC MEDLINE=89025644; PubMed=3178743;
RA Tromp G., Kuivaniemi H., Stacey A., Shikata H., Baldwin C.T.,
RA Jaenisch R., Prockup D.J.;
RT "Structure of a full-length cDNA clone for the prepro alpha 1(I)
RT chain of human type I procollagen.";
RL Biochem. J. 253:919-922(1988).
RN [2]
RP SEQUENCE OF 1-181 FROM N.A.
RC MEDLINE=84270697; PubMed=6462220;
RA Chu M.-L., de Wet W.J., Bernard M.P., Ding J.-F., Morabito M.,
RA Myers J., Williams C., Ramirez F.;
RT "Human pro alpha 1(I) collagen gene structure: reveals evolutionary
RT conservation of a pattern of introns and exons.";
RL Nature 310:337-340(1984).
RN [3]
RP SEQUENCE OF 162-301.
RC TISSUE=Skin;
RA MEDLINE=71038625; PubMed=5529814;
RT Click E.M., Bornstein P.;
RT "Isolation and characterization of the cyanogen bromide peptides from
RT the alpha 1 and alpha 2 chains of human skin collagen.";
RL Biochemistry 9:4699-4706(1970).
RN [4]
RP SEQUENCE OF 263-268.
RC TISSUE=Skin;
RA MEDLINE=71001508; PubMed=4319110;
RT Morgan P.H., Jacobs H.G., Segrest J.P., Cunningham L.W.;
RT "A comparative study of glycopeptides derived from selected
RT vertebrate collagens. A possible role of the carbohydrate in fibril
RT formation.";
RL J. Biol. Chem. 245:5042-5048(1970).
RN [5]
RP SEQUENCE OF 425-1464 FROM N.A.
RC MEDLINE=84080385; PubMed=6689127;
RA Bernard M.P., Chu M.-L., Myers J.C., Ramirez F., Eikenberry E.F.,
RA Prockup D.J.;
RT "Nucleotide sequences of complementary deoxyribonucleic acids for the
RT pro alpha 1 chain of human type I procollagen. Statistical evaluation
RT of structures that are conserved during evolution.";
RL Biochemistry 22:5213-5223(1983).
RN [6]
RP SEQUENCE OF 1229-1454 FROM N.A.
RC TISSUE=Bone;
RA MEDLINE=88124208; PubMed=3340531;
RT Maekelae J.K., Raassina M., Virta A., Vuorio E.;
RT "Human pro alpha 1(I) collagen: cDNA sequence for the C-propeptide
RT domain.";
RL Nucleic Acids Res. 16:349-349(1988).
RN [7]
RP SEQUENCE OF 1-34 FROM N.A.
RC MEDLINE=88097389; PubMed=3480516;
RA Bornstein P., McKay J., Morishima J.K., Devarayalu S., Gellinas R.E.;
RT "Regulatory elements in the first intron contribute to
RT transcriptional control of the human alpha 1(I) collagen gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8869-8873(1987).
RN [8]
RP SEQUENCE OF 1-34 FROM N.A.
RC MEDLINE=85130970; PubMed=2857713;
RA Chu M.-L., de Wet W.J., Bernard M.P., Ramirez F.;
RT "Fine structural analysis of the human pro-alpha 1 (I) collagen gene.
RT Promoter structure, Alu repeats, and polymorphic transcripts.";

```

RL J. Biol. Chem. 260:2315-2320(1985).
 RN [9]
 RP SEQUENCE OF 1-44 FROM N.A.
 RX MEDLINE-88033098; PubMed-2822714;
 RA Rossow C.M.S., Vergeer W.P., du Plooy S.J., Bernard M.P., Ramirez F.,
 de Wet W.J.;
 RT "DNA sequences in the first intron of the human pro-alpha 1(I)
 collagen gene enhance transcription.";
 RL J. Biol. Chem. 262:15151-15157(1987).
 RN [10]
 RP REVIEW ON VARIANTS.
 RX MEDLINE-91184577; PubMed-2010058;
 RA Kulvanliemi H., Tromp G., Prockop D.J.;
 RT "Mutations in collagen genes: causes of rare and some common diseases
 in humans.";
 RL FASEB J. 5:2052-2060(1991).
 RN [11]
 RP REVIEW ON VARIANTS.
 RX MEDLINE-97255959; PubMed-9101290;
 RA Kulvanliemi H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 associated collagen (type IX), and network-forming collagen (type X)
 cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315(1997).
 RN [12]
 RP REVIEW ON VARIANTS.
 RX MEDLINE-91374476; PubMed-1895312;
 RA Byers P.H., Wallis G.A., Willing M.C.;
 RT "Osteogenesis imperfecta: translation of mutation to phenotype.";
 RL J. Med. Genet. 28:433-442(1991).
 RN [13]
 RP REVIEW ON VARIANTS.
 RX MEDLINE-97169389; PubMed-9016532;
 RA Daigleish R.;
 RT "The human type I collagen mutation database.";
 RL Nucleic Acids Res. 25:181-187(1997).
 RN [14]
 RP VARIANT OI-II CYS-1166.
 RX MEDLINE-86287390; PubMed-3016737;
 RA Cohn D.H., Byers P.H., Steinmann B., Gellinas R.E.;
 RT "Lethal osteogenesis imperfecta resulting from a single nucleotide
 change in one human pro alpha 1(I) collagen allele.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6045-6047(1986).
 RN [15]
 RP VARIANT OI-II ARG-569.
 RX MEDLINE-87222293; PubMed-3108247;
 RA Bateman J.F., Chan D., Walkers I.D., Rogers J.G., Cole W.G.;
 RT "Lethal perinatal osteogenesis imperfecta due to the substitution of
 arginine for glycine at residue 391 of the alpha 1(I) chain of type I
 collagen.";
 RL J. Biol. Chem. 262:7021-7027(1987).
 RN [16]
 RP VARIANT OI-II CYS-926.
 RX MEDLINE-88033031; PubMed-3667599;
 RA Vogel B.E., Minor R.R., Freund M., Prockop D.J.;
 RT "A point mutation in a type I procollagen gene converts glycine 748
 of the alpha 1 chain to cysteine and destabilizes the triple helix in
 a lethal variant of osteogenesis imperfecta.";
 RL J. Biol. Chem. 262:14737-14744(1987).
 RN [17]
 RP VARIANT OI-II ARG-842.
 RX MEDLINE-86298828; PubMed-3403550;
 RA Bateman J.F., Lamande S.R., Dahl H.H., Chan D., Cole W.G.;
 RT "Substitution of arginine for glycine 664 in the collagen alpha 1(I)
 chain in lethal perinatal osteogenesis imperfecta. Demonstration of
 the peptide defect by in vitro expression of the mutant cDNA.";
 RL J. Biol. Chem. 263:11627-11630(1988).
 RN [18]
 RP VARIANT OI CYS-1195.
 RX MEDLINE-89218628; PubMed-3244312;
 RA Labhard M.E., Wirtz M.K., Pope F.M., Nicholls A.C., Hollister D.W.;
 RT "A cysteine for glycine substitution at position 1017 in an alpha
 1(I) chain of type I collagen in a patient with mild dominantly

RT inherited osteogenesis imperfecta.";
 RL Mol. Biol. Med. 5:197-207(1988).
 RN [19]
 RP VARIANT OI-II VAL-434.
 RX MEDLINE-89255493; PubMed-2470760;
 RA Patterson E., Smiley E., Bonadio J.;
 RT "RNA sequence analysis of a perinatal lethal osteogenesis imperfecta
 mutation.";
 RL J. Biol. Chem. 264:10083-10087(1989).
 RN [20]
 RP VARIANT OI-IV SER-1010.
 RX MEDLINE-89308591; PubMed-2745420;
 RA Marini J.C., Grange D.K., Gottesman G.S., Lewis M.B., Koepf D.A.;
 RT "Osteogenesis imperfecta type IV. Detection of a point mutation in
 one alpha 1(I) collagen allele (COL1A1) by RNA/RNA hybrid analysis.";
 RL J. Biol. Chem. 264:11893-11900(1989).
 RN [21]
 RP VARIANTS OI-II ALA-1106; VAL-1151; ARG-1154 AND VAL-1184.
 RX MEDLINE-89380165; PubMed-2777764;
 RA Lamande S.R., Dahl H.H., Cole W.G., Bateman J.F.;
 RT "Characterization of point mutations in the collagen COL1A1 and
 COL1A2 genes causing lethal perinatal osteogenesis imperfecta.";
 RL J. Biol. Chem. 264:15809-15812(1989).
 RN [22]
 RP VARIANT OI SER-1022.
 RX MEDLINE-90062068; PubMed-2511192;
 RA Pack M., Constantinou C.D., Kalia K., Nielsen K.B., Prockop D.J.;
 RT "Substitution of serine for alpha 1(I)-glycine 844 in a severe
 variant of osteogenesis imperfecta minimally destabilizes the triple
 helix of type I procollagen. The effects of glycine substitutions on
 thermal stability are either position of amino acid specific.";
 RL J. Biol. Chem. 264:19694-19699(1989).
 RN [23]
 RP VARIANT OI-II CYS-1082.
 RX MEDLINE-89109573; PubMed-2913053;
 RA Constantinou C.D., Nielsen K.B., Prockop D.J.;
 RT "A lethal variant of osteogenesis imperfecta has a single base
 mutation that substitutes cysteine for glycine 904 of the alpha 1(I)
 chain of type I procollagen. The asymptomatic mother has an
 unidentified mutation producing an overmodified and unstable type I
 procollagen.";
 RL J. Clin. Invest. 83:574-584(1989).
 RN [24]
 RP VARIANTS OI CYS-272; CYS-704 AND CYS-896.
 RX MEDLINE-90009313; PubMed-2794057;
 RA Starman B.J., Eyre D., Charbonneau H., Harrylock M., Weis M.A.,
 Weiss L., Graham J.M., Byers P.H.;
 RT "Osteogenesis imperfecta. The position of substitution for glycine by
 cysteine in the triple helical domain of the pro alpha 1(I) chains of
 type I collagen determines the clinical phenotype.";
 RL J. Clin. Invest. 84:1206-1214(1989).
 RN [25]
 RP VARIANT OI-II CYS-422.
 QY Query Match 100.0%; Score 80; DB 1; Length 1464;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 947 GTPGPGIAGRGV 15
 1 GTPGPGIAGRGV 15
 |||||
 RESULT 6
 CA12 HUMAN STANDARD; PRT; 1418 AA.
 ID CA12 HUMAN
 AC P02458;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Collagen alpha 1(I) chain precursor (Contains: Chondrocalcin).
 GN COL2A1.
 OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90067946; PubMed=2587267;
 RA Su M.W., Lee B., Ramirez F., Machado M., Horton W.;
 RT "Nucleotide sequence of the full length cDNA encoding for human type
 II procollagen.";
 RL Nucleic Acids Res. 17:9473-9473(1989).
 RN [2]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=87031574; PubMed=3021582;
 RA Nunez A.M., Kohno K., Martin G.R., Yamada Y.;
 RT "Promoter region of the human pro-alpha 1(II)-collagen gene.";
 RL Gene 44:11-16(1986).
 RN [3]
 RP SEQUENCE OF 432-1145 FROM N.A.
 RA Ramirez F.;
 RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 963-1418 FROM N.A.
 RX MEDLINE=85190534; PubMed=3857598;
 RA Cheah K.S.E., Stoker N.G., Griffin J.R., Grosveld F.G., Solomon E.;
 RT "Identification and characterization of the human type II collagen
 gene (COL2A1).";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2555-2559(1985).
 RN [5]
 RP SEQUENCE OF 1120-1398 FROM N.A.
 RX MEDLINE=85306861; PubMed=3840017;
 RA Elima K., Mäkelä J.K., Vuorio T., Kauppinen S., Knowles J.,
 RA Vuorio E.;
 RT "Construction and identification of a cDNA clone for human type II
 procollagen mRNA.";
 RL Biochem. J. 229:183-188(1985).
 RN [6]
 RP SEQUENCE OF 1106-1418 FROM N.A.
 RX MEDLINE=88067771; PubMed=2825137;
 RA Elima K., Vuorio T., Vuorio E.;
 RT "Determination of the single polyadenylation site of the human pro
 alpha 1(II) collagen gene.";
 RL Nucleic Acids Res. 15:9499-9504(1987).
 RN [7]
 RP SEQUENCE OF 1227-1289 FROM N.A.
 RX MEDLINE=86104139; PubMed=3002437;
 RA Nunez A.M., Francosano C., Young M.F., Martin G.R., Yamada Y.;
 RT "Isolation and partial characterization of genomic clones coding for
 a human pro-alpha 1 (II) collagen chain and demonstration of
 restriction fragment length polymorphism at the 3' end of the gene.";
 RL Biochemistry 24:6343-6348(1985).
 RN [8]
 RP SEQUENCE OF 1176-1226 FROM N.A.
 RX MEDLINE=84118798; PubMed=6320112;
 RA Strom C.M., Upholt W.B.;
 RT "Isolation and characterization of genomic clones corresponding to
 the human type II procollagen gene.";
 RL Nucleic Acids Res. 12:1025-1038(1984).
 RN [9]
 RP SEQUENCE OF 35-167 FROM N.A.
 RX MEDLINE=89233138; PubMed=2714801;
 RA Su M.W., Benson-Chanda V., Vissing H., Ramirez F.;
 RT "Organization of the exons coding for pro alpha 1(II) collagen N-
 propeptide confirms a distinct evolutionary history of this domain of
 the fibrillar collagen genes.";
 RL Genomics 4:438-441(1989).
 RN [10]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=91184577; PubMed=2010058;
 RA Kuivaniemi H., Tromp G., Prockop D.J.;
 RT "Mutations in collagen genes: causes of rare and some common diseases
 in humans.";
 RL FASEB J. 5:2052-2060(1991).
 RN [11]

RP REVIEW ON VARIANTS.
 RX MEDLINE=97255959; PubMed=9101290;
 RA Kuivaniemi H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 associated collagen (type IX), and network-forming collagen (type X)
 cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315(1997).
 RN [12]
 RP VARIANT SER-1074.
 RX MEDLINE=90036909; PubMed=2572591;
 RA Vissing H., D'Alessio M., Lee B., Ramirez F., Godfrey M.,
 RA Hollister D.W.;
 RT "Glycine to serine substitution in the triple helical domain of pro-
 alpha 1 (II) collagen results in a lethal perinatal form of short-
 limbed dwarfism.";
 RL J. Biol. Chem. 264:18265-18267(1989).
 RN [13]
 RP VARIANT SEDC 1095-GLY--TYR-1330 DEL.
 RX MEDLINE=89266907; PubMed=2543071;
 RA Lee B., Vissing H., Ramirez F., Rogers D., Rimoin D.;
 RT "Identification of the molecular defect in a family with
 spondyloepiphyseal dysplasia.";
 RL Science 244:978-980(1989).
 RN [14]
 RP VARIANT OSTEOARTHRITIS CYS-650.
 RX MEDLINE=90370826; PubMed=1975693;
 RA Ala-Kokko L., Baldwin C.T., Moskowitz R.W., Prockop D.J.;
 RT "Single base mutation in the type II procollagen gene (COL2A1) as a
 cause of primary osteoarthritis associated with a mild
 chondrodysplasia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6565-6568(1990).
 RN [15]
 RP VARIANT OI-IV VAL-717.
 RX MEDLINE=91291136; PubMed=2064612;
 RA Bateman J.F., Hannagan M., Chan D., Cole W.G.;
 RT "Characterization of a type I collagen alpha 2(I) glycine-586 to
 valine substitution in osteogenesis imperfecta type IV. Detection of
 the mutation and prenatal diagnosis by a chemical cleavage method.";
 RL Biochem. J. 276:765-770(1991).
 RN [16]
 RP VARIANT OSTEOARTHRITIS CYS-650.
 RX MEDLINE=91086471; PubMed=1985108;
 RA Eyre D.R., Weis M.A., Moskowitz R.W.;
 RT "Cartilage expression of a type II collagen mutation in an inherited
 form of osteoarthritis associated with a mild chondrodysplasia.";
 RL J. Clin. Invest. 87:357-361(1991).
 RN [17]
 RP VARIANT HYPOCHONDROGENESIS GLU-984.
 RX MEDLINE=93034548; PubMed=1429602;
 RA Bogaert R., Tiller G.E., Wiles M.A., Gruber H.E., Rimoin D.L.,
 RA Cohn D.H., Eyre D.R.;
 RT "An amino acid substitution (Gly983-->Glu) in the collagen alpha
 1(II) chain produces hypochondrogenesis.";
 RL J. Biol. Chem. 267:22522-22526(1992).
 RN [18]
 RP VARIANT HYPOCHONDROGENESIS SER-705.
 RX MEDLINE=92262484; PubMed=1374906;
 RA Horton W.A., Machado M.A., Ellard J., Campbell D., Bartley J.,
 RA Ramirez F., Vitale E., Lee B.;
 RT "Characterization of a type II collagen gene (COL2A1) mutation
 identified in cultured chondrocytes from human hypochondrogenesis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:4583-4587(1992).
 RN [19]
 RP VARIANT WS-II ASP-198.
 RX MEDLINE=93304428; PubMed=8317498;
 RA Koerkoe J., Ritsvanem P., Haataja L., Kaaerlahti H.,
 RA Kuivaniemi H., Prockop D.J., Ala-Kokko L.;
 RT "Mutation in type II procollagen (COL2A1) that substitutes aspartate
 for glycine alpha 1-67 and that causes cataracts and retinal
 detachment: evidence for molecular heterogeneity in the Wagner
 syndrome and the Stickler syndrome (arthro-ophthalmopathy).";
 RL Am. J. Hum. Genet. 53:55-61(1993).
 RN [20]

RA VARIANT SEMD CVS-840.
 RA Tyller G.E., Wells M.A., Lachman R.S., Cohn D.H., Rimoin D.L.,
 RA Eyre D.R.;
 RT "A dominant mutation in the type II collagen gene (COL2A1) produces
 RT spondyloepiphyseal dysplasia (SEMD), Strudwick type.";
 RL Am. J. Hum. Genet. 53:A209-A209(1993).
 RN [21]
 RP VARIANT OSTEOARTHRITIS CVS-650.
 RP MEDLINE-93282819; PubMed-8507190;
 RA Holderbaum D., Malesud C.J., Moskowitz R.W., Haq T.M.;
 RT "Human cartilage from late stage familial osteoarthritis transcribes
 RT type II collagen mRNA encoding a cysteine in position 519.";
 RL Biochem. Biophys. Res. Commun. 192:1169-1174(1993).
 RN [22]
 RP VARIANT SEMD ARG-285.
 RP MEDLINE-93252400; PubMed-8486375;
 RA Vilkula M., Rittvanlehti P., Vuorio A.F., Kaitila I., Ala-Kokko L.,
 RA Peltonen L.;
 RT "A mutation in the amino-terminal end of the triple helix of type II
 RT collagen causing severe osteochondrodysplasia.";
 RL Genomics 16:282-285(1993).
 RN [23]
 RP VARIANT SEDC CVS-206.
 RP MEDLINE-94063862; PubMed-8244341;
 RA Williams C.J., Considine E.L., Knowlton R.G., Reginato A., Neumann G.,
 RA Harrison D., Buxton P., Jimenez S.A., Prockop D.J.;
 RT "Spondyloepiphyseal dysplasia and precocious osteoarthritis in a
 RT family with an Arg75->Cys mutation in the procollagen type II gene
 RT (COL2A1).";
 RL Hum. Genet. 92:499-505(1993).
 RN [24]
 RP VARIANT SEDC CVS-920.
 RP MEDLINE-93315508; PubMed-8325895;
 RA Chan D., Taylor T.K.F., Cole W.G.;
 RT "Characterization of an arginine 789 to cysteine substitution in
 RT alpha 1 (II) collagen chains of a patient with spondyloepiphyseal
 RT dysplasia.";
 RL J. Biol. Chem. 268:15238-15245(1993).
 RN [25]
 RP VARIANT SEDC SER-1128.
 RP MEDLINE-93140139; PubMed-8423604;
 RA Cole W.G., Hall R.K., Rogers J.G.;
 RT "The clinical features of spondyloepiphyseal dysplasia congenita
 RT resulting from the substitution of glycine 997 by serine in the alpha
 RT 1(II) chain of type II collagen.";
 RL J. Med. Genet. 30:27-35(1993).

Query Match 88.88; Score 71; DB 1; Length 1418;
 Best Local Similarity 80.08; Pred. NO. 0.0049;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTGPGGAGGAGV 15
 Db 900 GPGPGGAGGAGV 914
 1 |||||:||||:|

RESULT 7
 CA12_MOUSE STANDARD; PRT: 1459 AA.
 AC P28481;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Collagen alpha 1(II) chain precursor [Contains: Chondrocalcin].
 GN COL2A1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RP MEDLINE-91358489; PubMed-1885613;
 RA Metsaanta M., Toman D., de Crombrughe B., Vuorio E.;

Query Match	Best Local Similarity	88.8%	Score 71; DB 1; Length 1459;
Matches	12; Conservative	80.0%;	Pred. No. 0.0051;
		2; Mismatches	1; Indels
			Gaps
QY	1 GTPGPGIAGRGVY 15		
DB	941 GPGPGIAGRGVY 955		
RESULT 8			
CA24_ASCSU	STANDARD;	PRT; 1763 AA.	
ID	CA24_ASCSU		

AC P27393; 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Collagen alpha 2(IV) chain precursor.
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
 OC Ascarididae; Ascaris.
 NCBI_TaxID=6253;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS I AND II).
 RX MEDLINE=91340768; PubMed=1714907;
 RA Pettitt J., Kingston I.B.,
 RT "The complete primary structure of a nematode alpha 2(IV) collagen
 and the partial structural organization of its gene."
 RL J. Biol. Chem. 266:16149-16156(1991).
 CC -I- FUNCTION: COLLAGEN TYPE IV IS SPECIFIC FOR BASEMENT MEMBRANS.
 CC -I- SUBUNIT: TRIMERS OF TWO ALPHA 1(IV) AND ONE ALPHA 2(IV) CHAIN.
 CC TYPE IV COLLAGEN FORMS A MESH-LIKE NETWORK LINKED THROUGH
 CC INTERMOLECULAR INTERACTIONS BETWEEN 7S DOMAINS AND BETWEEN NCI
 CC DOMAINS.
 CC -I- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing: Named isoforms-2;
 CC Name-I;
 CC IsoId=P27393-1; Sequence=Displayed;
 CC Name-II;
 CC IsoId=P27393-2; Sequence=VSP_001159;
 CC -I- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 CC TRIPLE-HELICAL 7S DOMAIN.
 CC -I- PWM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -I- PIR: Type IV collagens contain numerous cysteine residues which
 CC are involved in inter- and intramolecular disulfide bonding. 12 of
 CC these, located in the NC1 domain, are conserved in all known type
 CC IV collagens.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M67507; AAA18014.1; -;
 CC PIR: S16366; S16366.
 CC InterPro: IPR000087; Collagen.
 CC InterPro: IPR001442; ProcollagenC4.
 CC Pfam: PF01413; C4; 2.
 CC Pfam: PF01391; Collagen; 25.
 CC ProDom: PD000007; C1g helix; 6.
 CC ProDom: PD003923; ProcollagenC4; 1.
 CC SMART: SM00111; C4; 2.
 CC Hydroxylation: Connective tissue; Basement membrane; Repeat; Collagen;
 CC Alternative splicing: Glycoprotein; Signal.
 CC SIGNAL 1 26
 CC CHAIN 27 1763 COLLAGEN ALPHA 2(IV) CHAIN.
 CC DOMAIN 27 42
 CC DOMAIN 43 1529
 CC DOMAIN 1530 1763 TRIPLE-HELICAL REGION.
 CC DISULFID 1548 1637 NONHELMIC REGION (NC1).
 CC DISULFID 1581 1634 OR 1634 (BY SIMILARITY).
 CC DISULFID 1593 1599 OR 1637 (BY SIMILARITY).
 CC DISULFID 1656 1752 BY SIMILARITY.
 CC DISULFID 1690 1749 OR 1749 (BY SIMILARITY).
 CC DISULFID 1702 1709 OR 1752 (BY SIMILARITY).
 CC CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 249 249 O-LINKED (XYL. . .) (GLYCOSAMINOGLYCAN)
 CC (IN ISOFORM II) (POTENTIAL).
 CC VARSPLIC 230 266 GEOGPRGQPGPVPSTGACGTTTICPGAGMGK ->

FT FT GDICGAPGPGPPREFRTSGSIVGRGHSQDKGVK (in
 FT FT Isoform II).
 FT FT /FTId=VSP_001159.
 SQ SEQUENCE 1763 AA; 168526 MW; 304F52B8C06AA80D CRC64;
 Query Match 77.5%; Score 62; DB 1; Length 1763;
 Best Local Similarity 66.7%; Pred. No. 0.14;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GTPGPGIAGGQGV 15
 Db 82 GPPGPGIKGDRGIT 96
 RESULT 9
 CA29_HUMAN STANDARD; PRT; 689 AA.
 AC Q14055;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Collagen alpha 2(IX) chain precursor.
 GN COL9A2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ForeSkin;
 RX MEDLINE=98370844; PubMed=9707347;
 RA Pihlajamaa T., Vuorio M.M., Annunen S., Peraelae M., Prockop D.J.,
 RA Ala-Kokko L.;
 RT "Human COL9A1 and COL9A2 genes. Two genes of 90 and 15 kb code for
 RT similar polypeptides of the same collagen molecule."
 RT Matrix Biol. 17:237-241(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Donnelly S.;
 RL submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 50-668 FROM N.A.
 RC TISSUE=Cartilage;
 RX MEDLINE=9320262; PubMed=8454052;
 RA Peraelae M., Hanninen M., Hasbicka J., Elima K., Vuorio E.;
 RT "Molecular cloning of the human alpha 2(IX) collagen cDNA and
 RT assignment of the human COL9A2 gene to chromosome 1.";
 RL FEBS Lett. 319:177-180(1993).
 RN [4]
 RP VARIANT ID: TRP-326, AND VARIANT ARG-326.
 RX MEDLINE=99340300; PubMed=10411504;
 RA Annunen S., Paasilta P., Lohiniva J., Peraelae M., Pihlajamaa T.,
 RA Karppinen J., Tervonen O., Kroege H., Laehte S., Vanharanta H.,
 RA Ryhanen L., Goering H.H., Ott J., Prockop D.J., Ala-Kokko L.;
 RT "An allele of COL9A2 associated with intervertebral disc disease.";
 RL Science 285:409-412(1999).
 CC -I- FUNCTION: Structural component of hyaline cartilage and vitreous
 CC of the eye.
 CC -I- SUBUNIT: Heterotrimer of an alpha 1(IX), an alpha 2(IX) and an
 CC alpha 3(IX) chain.
 CC -I- PWM: Covalently linked to the telopeptides of type II collagen by
 CC lysine-derived cross-links.
 CC -I- PWM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -I- DISEASE: Defects in COL9A2 are a cause of multiple epiphyseal
 CC dysplasia type 2 (MED2) [MIM:600204]. MED2 is characterized by
 CC flattened, irregular epiphyses in most joints, particularly the
 CC knees. In childhood, typical features include waddling gait and
 CC knee pain/stiffness. Inheritance is autosomal dominant.
 CC -I- DISEASE: Defects in COL9A2 are a cause of susceptibility to
 CC intervertebral disc disease (IDD) [MIM:603932], one of the
 CC most common musculo-skeletal disorders.
 CC -I- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH

----- INTERRUPTED HELICES (FACIT) FAMILY -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF019406; AAC33512.1; -
DR EMBL; AL050341; CAB81611.1; -
DR EMBL; M95610; AAA80977.1; -
DR Genew; HGNC:2218; COL9A2.
DR MIM; 120260; -
DR MIM; 600204; -
DR MIM; 603932; -
DR GO; GO:0005594; C:collagen type IX; TAS.
DR GO; GO:0005202; F:collagen; TAS.
DR GO; GO:0001501; F:skeletal development; TAS.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF01391; Collagen; 9.
DR Prodom; PD000007; Ctg helix; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cartilage; Collagen; Signal; Disease mutation; Polymorphism.
FT SIGNAL 1 23
FT CHAIN 1 23
FT DOMAIN 24 689 COLLAGEN ALPHA 2(IX) CHAIN.
FT DOMAIN 27 519 TRIPLE-HELICAL REGION 3 (COL3).
FT DOMAIN 520 549 NONHELIICAL REGION 3 (NC3).
FT DOMAIN 550 632 TRIPLE-HELICAL REGION 2 (COL2).
FT DOMAIN 633 634 NONHELIICAL REGION 2 (NC2).
FT DOMAIN 635 664 TRIPLE-HELICAL REGION 1 (COL1).
FT DOMAIN 665 689 NONHELIICAL REGION 1 (NC1).
FT VARIANT 326 326 Q -> R.
FT /FTID=VAR_012659.
FT VARIANT 326 326 Q -> W (In IDD; requires 2 nucleotide substitutions).
FT /FTID=VAR_012658.
SQ SEQUENCE 689 AA; 65131 MW; EB6106E02FEFA862 CRC64;

Query Match 73.8%; Score 59; DB 1; Length 689;
Best Local Similarity 71.4%; Pred. No. 0.16;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DQ 1 GTPGPGIAGRGV 14
| | | | | : | | | | |
Db 493 GPPGPRLAGNRGV 506

RESULT 10
CA21_CHICK STANDARD: PRJ: 1362 AA.
AC P02457; P87491; P87492; Q90758; Q90792; Q90795; Q90797; Q92014;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Collagen alpha 2(I) chain precursor (Fragments).
GN COL1A2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE OF 1-245; 262-448 AND 466-1362 FROM N.A.
RX MEDLINE=86185168; PubMed=3868961;
RA Boedtker H., Finer M., Aho S.;
RT "The structure of the chicken alpha 2 collagen gene."
RL Ann. N.Y. Acad. Sci. 460:85-116(1985).
RN [2]
RP SEQUENCE OF 1-89 FROM N.A.
RX MEDLINE=83246518; PubMed=6135195;
RA Tate V.E., Finer M.H., Boedtker H., Doty P.;

RT "Chick pro alpha 2 (I) collagen gene: exon location and coding potential for the prepropeptide."
RL Nucleic Acids Res. 11:91-104(1983).
RN [3]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE=82060240; PubMed=6946474;
RA Vogel J.G., Ohkubo H., Sobel M.E., Yamada Y., Pastan I., de Crombrughe B.;
RT "Structure of the promoter for chicken alpha 2 type I collagen gene." proc. Natl. Acad. Sci. U.S.A. 78:5334-5338(1981).
RN [4]
RP SEQUENCE OF 1-33 FROM N.A.
RX MEDLINE=84297217; PubMed=6473103;
RA Ano S., Tate V.E., Boedtker H.;
RT "Location of the 11 bp exon in the chicken pro alpha 2(I) collagen gene." Nucleic Acids Res. 12:6117-6125(1984).
RN [5]
RP SEQUENCE OF 1-79 FROM N.A.
RX MEDLINE=88056316; PubMed=3678834;
RA Finer M.H., Boedtker H., Doty P.;
RT "Construction and characterization of cDNA clones encoding the 5' end of the chicken pro alpha 1(I) collagen mRNA." Gene 56:71-78(1987).
RN [6]
RP SEQUENCE OF 78-92.
RC TISSUE=Skin;
RX MEDLINE=7115216; PubMed=5544653;
RA Higberger J.H., Kang A.H., Gross J.;
RT "Comparative studies on the amino acid sequence of the alpha 2-CB2 peptides from chick and rat skin collagens." Biochemistry 10:610-616(1971).
RN [7]
RP SEQUENCE OF 74-91; 263-448 AND 1088-1169 FROM N.A.
RX MEDLINE=82058081; PubMed=6272119;
RA Wozney J., Hanahan D., Tate V.E., Boedtker H., Doty P.;
RT "Structure of the pro alpha 2 (I) collagen gene." Nature 294:129-135(1981).
RN [8]
RP SEQUENCE OF 78-92.
RC TISSUE=Skin;
RX MEDLINE=70131186; PubMed=4313735;
RA Kang A.H., Gross J.;
RT "Amino acid sequence of cyanogen bromide peptides from the amino-terminal region of chick skilcollagen." Biochemistry 9:796-804(1970).
RN [9]
RP SEQUENCE OF 78-92 AND 415-448.
RC TISSUE=Skin;
RX MEDLINE=69285369; PubMed=5809220;
RA Kang A.H., Igarashi S., Gross J.;
RT "Characterization of the cyanogen bromide peptides from the alpha-2 chain of chick skin collagen." Biochemistry 8:3200-3204(1969).
RN [10]
RP SEQUENCE OF 78-92 AND 415-448.
RC TISSUE=Bone;
RX MEDLINE=69206882; PubMed=5785233;
RA Lane J.M., Miller E.J.;
RT "Isolation and characterization of the peptides derived from the alpha 2 chain of chick bone collagen after cyanogen bromide cleavage." Biochemistry 8:2134-2139(1969).
RN [11]
RP SEQUENCE OF 566-587 FROM N.A.
RX MEDLINE=79074829; PubMed=364479;
RA Lehrach H., Frischau A.-M., Hanahan D., Wozney J., Fuller F., Crikvenjakov R., Boedtker H., Doty P.;
RT "Construction and characterization of a 2.5-kilobase procollagen clone." Proc. Natl. Acad. Sci. U.S.A. 75:5417-5421(1978).
RN [12]
RP SEQUENCE OF 902-1362 FROM N.A.

RX MEDLINE=81160715; PubMed=6927845;
 RA Fuller F., Boedtker H.;
 RT "Sequence determination and analysis of the 3' region of chicken pro-
 alpha 1(I) and pro-alpha 2(I) collagen messenger ribonucleic acids
 including the carboxy-terminal propeptide sequences.";
 RL Biochemistry 20:996-1006(1981).
 RN [13]
 RP SEQUENCE OF 998-1169 AND 1234-1362 FROM N.A.
 RX MEDLINE=81264246; PubMed=6267043;
 RA Dickson L.A., Ninomiya Y., Bernard M.P., Pesciotta D.M., Parsons J.,
 Green G., Eikenberry E.F., de Crombrughe B., Vogeli G., Pastan I.,
 Fietzek P.P., Olsen B.R.;
 RT "The exon/intron structure of the 3'-region of the pro alpha 2(I)
 collagen gene";
 RL J. Biol. Chem. 256:8407-8415(1981).
 RN [14]
 RP SEQUENCE OF 932-954 AND 968-980 FROM N.A.
 RX MEDLINE=81064671; PubMed=6159982;
 RA Avvedimento V.E., Vogeli G., Yamada Y., Maizel J.V. Jr., Pastan I.,
 de Crombrughe B.;
 RT "Correlation between splicing sites within an intron and their
 sequence complementarity with U1 RNA.";
 RL Cell 21:689-696(1980).
 RN [15]
 RP SEQUENCE OF 126-161, 467-517 AND 926-954 FROM N.A.
 RX MEDLINE=8112157; PubMed=7460017;
 RA Yamada Y., Avvedimento V.E., Mudry M., Ohkubo H., Vogeli G.,
 Imani M., Pastan I., de Crombrughe B.;
 RT "The collagen gene: evidence for its evolutionary assembly by
 amplification of a DNA segment containing an exon of 54 bp.";
 RL Cell 22:887-892(1980).
 CC -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
 CC (FIBRILLAR FORMING COLLAGEN).
 CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
 CC -1- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
 CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
 CC HYDROXYAPATITE.
 CC -1- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M25963; AAA69960.1; JOINED.
 DR EMBL: M25956; AAA69960.1; JOINED.
 DR EMBL: M25959; AAA69960.1; JOINED.
 DR EMBL: M25961; AAA69960.1; JOINED.
 DR EMBL: M25962; AAA69960.1; JOINED.
 DR EMBL: M25965; AAA69961.1; JOINED.
 DR EMBL: M25964; AAA69961.1; JOINED.
 DR EMBL: M25984; AAA69962.1; JOINED.
 DR EMBL: M25957; AAA69962.1; JOINED.
 DR EMBL: M25966; AAA69962.1; JOINED.
 DR EMBL: M25967; AAA69962.1; JOINED.
 DR EMBL: M25969; AAA69962.1; JOINED.
 DR EMBL: M25970; AAA69962.1; JOINED.
 DR EMBL: M25971; AAA69962.1; JOINED.
 DR EMBL: M25972; AAA69962.1; JOINED.
 DR EMBL: M25973; AAA69962.1; JOINED.
 DR EMBL: M25974; AAA69962.1; JOINED.
 DR EMBL: M25976; AAA69962.1; JOINED.
 DR EMBL: M25977; AAA69962.1; JOINED.
 DR EMBL: M25978; AAA69962.1; JOINED.
 DR EMBL: M25979; AAA69962.1; JOINED.
 DR EMBL: M25980; AAA69962.1; JOINED.
 DR EMBL: M25981; AAA69962.1; JOINED.
 DR EMBL: M25982; AAA69962.1; JOINED.
 DR EMBL: M25983; AAA69962.1; JOINED.

DR EMBL: J00826; AAA51611.1; JOINED.
 DR EMBL: J00821; AAA51611.1; JOINED.
 DR EMBL: K00792; AAA51611.1; JOINED.
 DR EMBL: J00830; AAA51613.1; JOINED.
 DR EMBL: J00829; AAA51613.1; JOINED.
 DR EMBL: J00837; AAA51614.1; JOINED.
 DR EMBL: J00812; AAA51615.1; JOINED.
 DR EMBL: J00811; AAA51615.1; JOINED.
 DR EMBL: J00814; AAA51615.1; JOINED.
 DR EMBL: J00815; AAA51615.1; JOINED.
 DR EMBL: X02557; CAA26493.1; JOINED.
 DR EMBL: K00794; NOT_ANNOTATED_CDS.
 DR EMBL: V00390; CAA23688.1; JOINED.
 DR EMBL: M17608; AAA48673.1; JOINED.
 DR EMBL: M10581; AAA48637.1; JOINED.
 DR EMBL: M10540; AAA48638.1; JOINED.
 DR EMBL: J00828; AAA51612.1; JOINED.
 DR EMBL: J00827; AAA51612.1; JOINED.
 DR EMBL: J00831; NOT_ANNOTATED_CDS.
 DR EMBL: J00832; NOT_ANNOTATED_CDS.
 DR EMBL: J00833; NOT_ANNOTATED_CDS.
 DR EMBL: J00822; NOT_ANNOTATED_CDS.
 DR PIR: I50173; I50173.
 DR PIR: I50206; CGCH2S.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib-collagen_C.
 DR Pfam: PF01391; Collagen; 17.
 DR Pfam: PF01410; COLFI; 1.
 DR ProDom: PD000007; Ctg_helix; 4.
 DR ProDom: PD002078; Fib-collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;

Query Match 72.5%; Score 58; DB 1; Length 1362;
 Best Local Similarity 66.7%; Pred. NO. 0.43;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GTPGPGIAGGQGVV 15
 Db 857 GTPGPGIAGGQGVV 871

RESULT 11
 CA25_HUMAN STANDARD; PRT; 1496 AA.
 ID CA25_HUMAN P05997;
 AC P05997;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Collagen alpha 2(V) chain precursor.
 GN COL5A2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 OX [1]
 RP SEQUENCE OF 1-463 FROM N.A.
 RX MEDLINE=89123368; PubMed=2914927;
 RA Woodbury D., Benson-Chanda V., Ramirez F.;
 RT "Amino-terminal propeptide of human pro-alpha 2(V) collagen conforms
 RT to the structural criteria of a fibrillar procollagen molecule.";
 RL J. Biol. Chem. 264:2735-2738(1989).
 RN [2]
 RP SEQUENCE OF 398-1496 FROM N.A.
 RX MEDLINE=87146331; PubMed=3029669;
 RA Weil D., Bernard M.P., Gargano S., Ramirez F.;
 RT "The pro alpha 2(V) collagen gene is evolutionarily related to the
 RT major fibrillar-forming collagens.";
 RL Nucleic Acids Res. 15:181-198(1987).
 RN [3]
 RP SEQUENCE OF 1227-1496 FROM N.A.
 RX MEDLINE=85289337; PubMed=2411731;
 RA Myers J.C., Loidl H.R., Seyer J.M., Dion A.S.;

RT "Complete primary structure of the human alpha 2 type V procollagen
RT COOH-terminal propeptide.";
RL J. Biol. Chem. 260:11216-11222(1985).
RN [4]
RP SEQUENCE OF 1449-1496 FROM N.A.
RX MEDLINE-69138450; PubMed-3224983;
RA Tsiipouras P., Schwartz R.C., Liddell A.C., Salkeid C.S., Weil D.,
RA Ramirez F.;
RT "Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2,
RT located on the long arm of human chromosome 2.";
RL Genomics 3:275-277(1988).
RN [5]
RP SEQUENCE OF 208-227.
RX MEDLINE-Piacenta; PubMed-1571108;
RA Mann K.;
RT "Isolation of the alpha 3-chain of human type V collagen and
RT characterization by partial sequencing.";
RL Biol. Chem. Hoppe-Seyler 373:69-75(1992).
RN [6]
RP SEQUENCE OF 288-297 AND 606-617.
RX TISSUE-Bone;
RA Medrati-Ameli M., Rousseau J.C., Klemm J.P., Champliand M.F.,
RA Bouillon M.M., Bernillon J., Wallach J.M., van der Rest M.;
RT "Diversity in the processing events at the N-terminus of type-V
RT collagen.";
RL Eur. J. Biochem. 221:987-995(1994).
RN [7]
RP DISEASE.
RX MEDLINE-98087576; PubMed-9425231;
RA Michalickova K., Susic M., Willing M.C., Wenstrup R.J., Cole W.G.;
RT "Mutations of the alpha2(V) chain of type V collagen impair matrix
RT assembly and produce Ehlers-Danlos syndrome type I.";
RL Hum. Mol. Genet. 7:249-255(1998).
RN [8]
RP VARIANT EDS-II ARG-960.
RX MEDLINE-98455031; PubMed-9783710;
RA Richards A.J., Martin S., Nicholls A.C., Harrison J.B., Pope F.M.,
RA Burrows N.P.;
RT "A single base mutation in COL5A2 causes Ehlers-Danlos syndrome type
RT II.";
RL J. Med. Genet. 35:846-848(1998).
RN [9]
RP FUNCTION: TYPE V COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
RP (FIBRILLAR FORMING COLLAGEN). IT IS A MINOR CONNECTIVE TISSUE
RP COMPONENT OF NEARLY UBQUITOUS DISTRIBUTION. TYPE V COLLAGEN BINDS
RP TO DNA, HEPARAN SULFATE, THROMBOSPONDIN, HEPARIN, AND INSULIN.
CC -1- SUBUNIT: TRIMERS OF TWO ALPHA 1(V) AND ONE ALPHA 2(V) CHAINS IN
CC MOST TISSUES AND TRIMERS OF ONE ALPHA 1(V), ONE ALPHA 2(V), AND
CC ONE ALPHA 3(V) CHAINS IN PLACENTA.
CC -1- PPM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- DISEASE: Defects in COL5A2 are a cause of Ehlers-Danlos syndrome
CC type I (EDS-I) [MIM:130000]; also known as Ehlers-Danlos syndrome
CC gravis. EDS-I is an autosomal dominant connective-tissue disorder
CC characterized by loose-jointedness and fragile, velvety,
CC stretchable, bruisable skin that heals with peculiar 'cigarette-
CC paper' scars. Inheritance is autosomal dominant.
CC -1- DISEASE: Defects in COL5A2 are a cause of Ehlers-Danlos syndrome
CC type II (EDS-II) [MIM:130010]; also known as Ehlers-Danlos
CC syndrome mitis. Inheritance is autosomal dominant.
CC -1- SIMILARITY: Contains 1 WMC domain.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
OR EMBL; J04478; AAAS1859.1; -;
OR EMBL; X04758; CAA28454.1; -;

DR	EMBL:	M1718;	AAAS2058.1; -.
DR	PIR:	A31427;	CGH02V.
DR	PDB:	1A9A;	18-NOV-98.
DR	Genew:	HGNC:2210;	COL5A2.
DR	MIM:	120190;	-.
DR	MIM:	130000;	-.
DR	MIM:	130010;	-.
DR	GO:	GO:0005588;	C:collagen type V; TAS.
DR	GO:	GO:0008202;	F:collagen; TAS.
DR	GO:	GO:0008151;	P:cell growth and/or maintenance; TAS.
DR	InterPro:	IPR000087;	Collagen.
DR	InterPro:	IPR000885;	Fib-collagen_C.
DR	InterPro:	IPR001007;	VWF_C.
DR	Pfam:	PF01410;	COLF1; 1.
DR	Pfam:	PF01391;	Collagen; 18.
DR	Pfam:	PF00093;	wvc; 1.
DR	ProdDom:	PD000007;	Ctg_helix; 5.
DR	ProdDom:	PD002078;	Fib_collagen_C; 1.
DR	SMART:	SMO0038;	COLF1; 1.
DR	SMART:	SMO0214;	VWC; 1.
DR	PROSITE:	PS01208;	VWFC_1; 1.
DR	PROSITE:	PS50184;	VWFC_2; 1.
KW	Extracellular matrix;	Connective tissue; Repeat;	Hydroxylation;
KW	Glycoprotein;	Collagen; Signal;	Ehlers-Danlos syndrome;
KW	Disease mutation;	3D-structure.	
FT	SIGNAL	1	26
FT	CHAIN	27	1226
FT	PROPEP	1227	1496
FT	DOMAIN	39	97
FT	MOD_RES	290	290
FT	MOD_RES	293	293
FT	MOD_RES	296	296
FT	MOD_RES	608	608
FT	MOD_RES	614	614
FT	VARIANT	960	960
FT	CONFLICT	292	292
FT	CONFLICT	1418	1418
FT	CONFLICT	1438	1438
FT	CONFLICT	1460	1460
FT	CONFLICT	1496	1496
SO	SEQUENCE	1496 AA;	144720 MW; 82827C17AB64AF5A CRC64;
<hr/>			
Query Match			
Best Local Similarity 72.5%; Score 58; DB 1; Length 1496;			
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0			
QY	1	GTGGPGAGGRGVV	15
Db	978	GPPGPAGTTGGRGIV	992
<hr/>			
RESULT 12			
ID	AI13_RAT	STANDARD;	PRT; 636 AA.
AC	P13941;	C70604;	
DT	01-JAN-1990	(Rel. 13,	Created)
DT	01-JUN-1994	(Rel. 29,	Last sequence update)
DT	16-OCT-2001	(Rel. 40,	Last annotation update)
DE	Collagen alpha 1(III) chain (fragment).		
GN	COL3A1.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=94114571;	PubMed=8286415;	
RA	Glumoff V., Maekelae J.K., Vuorio E.;		
RT	"Cloning of cDNA for rat pro alpha 1(III) collagen mRNA. Different		
RT	expression patterns of type I and type III collagen and fibronectin		
RT	genes in experimental granulation tissue."		
RL	Biochim. Biophys. Acta 1217:41-48(1994).		

RN [2]
 RP SEQUENCE OF 73-636 FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Fibroblast;
 RA Murtz T., Ellerstrom C., Lundmark C., Christenson C.;
 RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 308-482 FROM N.A.
 RX MEDLINE=88296083; PubMed=2456904;
 RA Frankel F.R., Hsu C.-Y.J., Meyers J.C., Lin E., Lytle C.R.,
 RA Komm B., Mohr K.;
 RT "Regulation of alpha 2(I), alpha 1(III), and alpha 2(V) collagen
 RT mRNAs by estradiol in the immature rat uterus.";
 RL DNA 7,347-354(1988).
 CC -1- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
 CC ALONG WITH TYPE I COLLAGEN.
 CC -1- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
 CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
 CC ALSO CROSS-LINKED VIA HYDROXYLINES.
 CC -1- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 DR EMBL: X70369; CAA49832.1; -;
 DR EMBL: AJ005395; CAA06510.1; -;
 DR EMBL: M21354; AAA40942.1; -;
 DR PIR: S41067; S41067.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLF1; 1.
 DR Pfam: PF01391; Collagen; 6.
 DR ProDom: PD000007; C1g_helix; 1.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SMO0038; COLF1; 1.
 DR PROSITE: PS01208; VWF_C_1; PARTIAL.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Collagen; Glycoprotein.
 FT CHAIN 1 375 COLLAGEN ALPHA 1(III) CHAIN.
 FT PROPEP 376 636 CARBOXYL-TERMINAL PROPEPTIDE.
 FT DOMAIN <1 368 TRIPLE-HELICAL REGION.
 FT DOMAIN 369 636 NONHELICAL REGION (C-TERMINAL).
 FT DISULFID 368 368 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 369 369 INTERCHAIN (BY SIMILARITY).
 FT CONFLICT 340 340 N -> D (IN REF. 2).
 FT CONFLICT 429 429 A -> G (IN REF. 2).
 SQ SEQUENCE 636 AA; 62332 MW; 61A48159F01D01EE CMC64;
 Query Match 71.2%; Score 57; DB 1; Length 636;
 Best Local Similarity 69.2%; Pred. No. 0.3;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE OF 1-242.
 RX MEDLINE=80026026; PubMed=488906;
 RA Fietzek P.P., Allmann H., Rautenberg J., Henkel W., Wächter E.,
 RA Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. I. The amino
 RT acid sequence of the amino terminal region of the alpha 1(III) chain
 RT (positions 1-222).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:809-820(1979).
 RN [2]
 RP SEQUENCE OF 243-422.
 RX MEDLINE=80026027; PubMed=488907;
 RA Dewes H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. II. The amino
 RT acid sequence of the cyanogen bromide peptide alpha 1(III)CB1,8,10,2
 RT (positions 223-402).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:821-832(1979).
 RN [3]
 RP SEQUENCE OF 423-571.
 RX MEDLINE=80026028; PubMed=488908;
 RA Bentz H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. III. The
 RT amino acid sequence of the cyanogen bromide peptide alpha 1(III)CB4
 RT (positions 403-551).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:833-840(1979).
 RN [4]
 RP SEQUENCE OF 572-808.
 RX MEDLINE=80026029; PubMed=488909;
 RA Lang H., Glanville R.W., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. IV. The amino
 RT acid sequence of the cyanogen bromide peptide alpha 1(III)CB5
 RT (positions 552-788).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:841-850(1979).
 RN [5]
 RP SEQUENCE OF 809-947.
 RX MEDLINE=80026030; PubMed=488910;
 RA Dewes H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. V. The amino
 RT acid sequence of the cyanogen bromide peptide alpha 1(III)CB9A
 RT (position 789-927).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:851-860(1979).
 RN [6]
 RP SEQUENCE OF 948-1049.
 RX MEDLINE=80026031; PubMed=488911;
 RA Allmann H., Fietzek P.P., Glanville R.W., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. VI. The amino
 RT acid sequence of the carboxyterminal cyanogen bromide peptide alpha
 RT 1(III)CB9B (positions 928-1028).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:861-868(1979).
 CC -1- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
 CC ALONG WITH TYPE I COLLAGEN.
 CC -1- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
 CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
 CC ALSO CROSS-LINKED VIA HYDROXYLINES.
 CC -1- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -----
 CC PIR: A02862; CGB07S.
 CC InterPro: IPR000087; Collagen.
 CC InterPro: IPR001007; VWF_C.
 CC Pfam: PF01391; Collagen; 17.
 CC ProDom: PD000007; C1g_helix; 3.
 CC PROSITE: PS01208; VWF_C_1; PARTIAL.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen.
 FT DOMAIN 1 14 NONHELICAL REGION (N-TERMINAL).
 FT DOMAIN 15 1040 TRIPLE-HELICAL REGION.
 FT DOMAIN 1041 1049 NONHELICAL REGION (C-TERMINAL).
 FT MOD_RES 95 95 HYDROXYLATION.

Query	Match	Best local	Similarity	Score	Length
QY	1 GTPGPCGACGRC 13	71.24;	69.24;	57;	1362;
Db	851 GTPGPCGACGRC 863	9;	2;	2;	0;
Matches	9;	Conservative	2;	Mismatches	2;
Indels	0;	Gaps	0;		

RC STRAIN=C57BL/6 X DBA; TISSUE=Embryo;
 RX MEDLINE=95011609; PubMed=7926795;
 RA Toman D., de Crombrughe B.;
 RT "The mouse type-III procollagen-encoding gene: genomic cloning and
 complete DNA sequence."
 RL Gene 147:161-168(1994).
 RN [2]
 RP SEQUENCE OF 1-488 FROM N.A.
 RX MEDLINE=88167858; PubMed=3443309;
 RA Wood L., Theriault N., Vogel G.;
 RT "Complete nucleotide sequence of the N-terminal domains of the murine
 alpha-1(I)-type-III collagen chain."
 RL Gene 61:225-230(1987).
 RN [3]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=85131189; PubMed=3972847;
 RA Lian G., Mudryj M., de Crombrughe B.;
 RT "Identification of the promoter and first exon of the mouse alpha 1
 (III) collagen gene."
 RL J. Biol. Chem. 260:3773-3777(1985).
 RN [4]
 RP SEQUENCE OF 810-1464 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
 RX MEDLINE=21083660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Akawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Flisbachman W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuhl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guinacich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [5]
 RP SEQUENCE OF 1442-1464 FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=91274355; PubMed=2054384;
 RA Metsaeranta M., Toman D., de Crombrughe B., Vuorio E.;
 RT "Specific hybridization probes for mouse type I, II, III and IX
 collagen mRNAs."
 RL Biochim. Biophys. Acta 1089:241-243(1991).
 CC -1- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
 CC ALONG WITH TYPE I COLLAGEN.
 CC -1- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
 CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
 CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
 CC -1- PTM: Proline residues at the third position of the tripeptide
 CC repeating unit (G-X-Y) are hydroxylated in some or all of the
 CC chains.
 CC -1- PTM: O-linked glycan consists of a Glc-Gal disaccharide bound to
 CC the oxygen atom of a post-translationally added hydroxyl group (By
 CC similarity).
 CC -1- SIMILARITY: Contains 1 WMEC domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X52046; CAA36279.1; -;
 DR EMBL: M18933; AAA37338.1; -;
 DR EMBL: K03037; -, NOT_ANNOTATED_CDS.
 DR EMBL: AK019446; BAB31724.1; -;
 DR EMBL: X57983; CAA41048.1; -;
 DR PIR: A27353; A27353.
 DR PIR: S59856; S59856.
 DR MGI: 88453; Col3a1.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; WFC_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR Prodom: PP000007; Ctg_helix; 1.
 DR Prodom: PP002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 DR SMART: SM00214; WFC; 1.
 DR PROSITE: PS01208; WFC_1; 1.
 DR PROSITE: PS0184; WFC_2; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen; Signal.
 FT SIGNAL 1 23 BY SIMILARITY.
 FT PROPEP 24 154 AMINO-TERMINAL PROPEPTIDE.
 FT CHAIN 155 1203 COLLAGEN ALPHA 1(III) CHAIN.
 FT PROPEP 1204 1464 CARBOXYL-TERMINAL PROPEPTIDE.
 FT DOMAIN 31 90 WFC.
 FT DOMAIN 155 169 NONHELICAL REGION (N-TERMINAL).
 FT DOMAIN 170 1195 TRIPLE-HELICAL REGION.
 FT DOMAIN 1196 1464 NONHELICAL REGION (C-TERMINAL).
 FT CARBOHD 262 262 O-LINKED (GAL. . .) (BY SIMILARITY).
 FT MOD_RES 262 262 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 283 283 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 859 859 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 976 976 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 1093 1093 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 1105 1105 HYDROXYLATION (BY SIMILARITY).
 FT DISULFD 1195 1195 INTERCHAIN (BY SIMILARITY).
 FT DISULFD 1196 1196 INTERCHAIN (BY SIMILARITY).
 SO SEQUENCE 1464 AA; 138944 MW; 2104EC27A886090B CRC64;

Query Match 71.2%; Score 57; DB 1; Length 1464;
 Best Local Similarity 69.2%; Pred. No. 0.65;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 GTPGPGIAGORG 13
 Db 851 GTPGPGVKGGRG 863
 Search completed: August 29, 2003, 18:27:05
 Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2003, 18:21:45; Search time 96 Seconds
(without alignments)
40.321 Million cell updates/sec

Title: US-09-935-417-1
Perfect score: 80
Sequence: 1 GTPGPGIAGORGVV 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	230	11	Q9R149
2	80	100.0	589	11	Q991L6
3	80	100.0	1453	11	Q63079
4	80	100.0	1461	4	Q76045
5	80	100.0	1464	4	Q8M473
6	75	93.8	809	13	Q93485
7	75	93.8	1449	13	Q910C0
8	74	92.5	1450	13	Q9Y1B4
9	72	90.0	678	13	Q93486
10	72	90.0	1458	13	Q910B9
11	71	88.8	113	11	Q8V172
12	71	88.8	347	6	Q9XT25
13	71	88.8	826	11	Q8K0M6
14	71	88.8	886	11	Q8CEP7
15	71	88.8	1160	4	Q14046
16	71	88.8	1418	6	Q28396

17	71	88.8	1418	13	Q9W7R9	Q9W7R9 cynops pyr
18	71	88.8	1419	11	Q63123	Q63123 rattus norv
19	71	88.8	1420	13	Q90W37	Q90W37 gallus gall
20	71	88.8	1442	11	Q62031	Q62031 mus musculu
21	71	88.8	1442	11	Q62033	Q62033 mus musculu
22	71	88.8	1459	11	Q62032	Q62032 mus musculu
23	71	88.8	1487	4	Q14047	Q14047 homo sapien
24	71	88.8	1487	6	Q77753	Q77753 canis famill
25	68	85.0	1486	13	Q91717	Q91717 xenopus lae
26	64	80.0	1445	13	Q93251	Q93251 rana catesb
27	64	80.0	1447	13	Q91B91	Q91B91 xenopus lae
28	63	78.8	441	4	Q96A83	Q96A83 homo sapien
29	60	75.0	100	4	Q9Y3P3	Q9Y3P3 homo sapien
30	60	75.0	441	4	Q96A84	Q96A84 homo sapien
31	58	72.5	290	5	Q26054	Q26054 paracentrot
32	58	72.5	1258	13	Q8AM11	Q8AM11 brachydantio
33	58	72.5	1347	4	Q960B3	Q960B3 homo sapien
34	58	72.5	1464	11	Q8BKX2	Q8BKX2 mus musculu
35	58	72.5	1497	11	Q61431	Q61431 mus musculu
36	57	71.2	310	13	Q90612	Q90612 gallus gall
37	57	71.2	832	4	Q961F7	Q961F7 homo sapien
38	57	71.2	998	11	Q8CFM4	Q8CFM4 mus musculu
39	57	71.2	1222	11	Q8K173	Q8K173 mus musculu
40	57	71.2	1414	5	Q26634	Q26634 strongyloce
41	57	71.2	1464	11	Q8BLM4	Q8BLM4 mus musculu
42	57	71.2	1860	4	Q81ZC6	Q81ZC6 homo sapien
43	56	70.0	690	13	Q8TGL8	Q8TGL8 brachydantio
44	56	70.0	1366	4	Q15177	Q15177 homo sapien
45	56	70.0	1621	4	Q9H4R9	Q9H4R9 homo sapien

ALIGNMENTS

RESULT 1	Q9R149	PRELIMINARY:	PRT:	230 AA.
AC	Q9R149:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DE	Pro-alpha-1 type 1 collagen (Fragment).			
OC	Cavia porcellus (Guinea pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.			
OX	NCBI_TaxID=10141;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Hartley; TISSUE=trachea;			
RA	Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,			
RA	Sekizawa K.;			
RT	"Guinea-pig pro-alpha-1 type 1 collagen expression in injured tracheal			
RT	epithelium.";			
RL	Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AF169346; AAD49346.1; -			
DR	InterPro: IPR000087; Collagen.			
DR	PFam: PF01391; Collagen; 4.			
KW	Collagen.			
FT	NON_TER	1	230	1
FT	NON_TER	230	230	1
SQ	SEQUENCE	230 AA:	20425 MW;	1A465F92779D9A71 CRC64;
Query Match	Best Local Similarity	100.0%;	Score 80;	DB 11; Length 230;
Matches	15; Conservative	0;	Pred. No. 0.00015;	
		0;	Mismatches	0; Indels
			Gaps	0;
OY	1 GTPGPGIAGORGVV 15			
Db				
	78 GTPGPGIAGORGVV 92			
RESULT 2	Q991L6			

ID 099L6 PRELIMINARY; PRT; 589 AA.
 AC 099L6;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, last annotation update)
 DE Hypothetical 58.8 kDa protein (Fragment).
 GN COL1A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC003198; AA03198.1;
 DR MGD; MGI:88467; Collal.
 DR InterPro; IPR000885; Fib_collagen.
 DR InterPro; IPR000885; Fib_collagen_C.
 DR Pfam; PF01410; COLFI; 1.
 DR Pfam; PF01391; Collagen; 6.
 DR ProDom; PD002078; Fib_collagen_C; 1.
 DR SMART; SM00038; COLFI; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 589 AA; 58805 MW; 81847495SE05CEP CRC64;

Query Match 100.0%; Score 80; DB 11; Length 589;
 Best Local Similarity 100.0%; Pred. No. 0.00039;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTPGPGIAGGRCVV 15
 Db 72 GTPGPGIAGGRCVV 86

RESULT 3

ID 063079 PRELIMINARY; PRT; 1453 AA.
 AC 063079;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-JUN-1998 (TREMblrel. 06, last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, last annotation update)
 DE Collagen alpha1 (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE OF 1-1092 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Tooth;
 RA Brandsten C., Lundmark C., Christerson C., Hammarstrom L., Wurtz T.;
 RT "Expression of Collagen alpha1(I) mRNA variants during Tooth and Bone
 Formation in the Rat."
 RT J. Dent. Res. 0:0-0(0).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Sprague-Dawley; TISSUE=Tooth;
 RA Wurtz T.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z78279; CAB01633.1;
 DR InterPro; IPR000887; Collagen.
 DR InterPro; IPR000885; Fib_collagen.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01410; COLFI; 1.
 DR Pfam; PF01391; Collagen; 18.
 DR ProDom; PD000007; Collagen; 3.
 DR ProDom; PD002078; Fib_collagen_C; 1.
 DR SMART; SM00038; COLFI; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS01208; WVC; 1.
 KW Collagen.
 FT NON_TER 1 1

SQ SEQUENCE 1453 AA; 137887 MW; E6896BDC19A4A1D8 CRC64;

Query Match 100.0%; Score 80; DB 11; Length 1453;
 Best Local Similarity 100.0%; Pred. No. 0.00096;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTPGPGIAGGRCVV 15
 Db 936 GTPGPGIAGGRCVV 950

RESULT 4

ID 076045 PRELIMINARY; PRT; 1461 AA.
 AC 076045;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 12, last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, last annotation update)
 DE Pro alpha 1(I) collagen.
 GN COL1A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chu M.L., de Wet W., Bernard M., Ramirez F.;
 RT "Fine structural analysis of the human pro-alpha 1 (I) collagen gene.
 promoter structure, AluI repeats, and polymorphic transcripts."
 RT J. Biol. Chem. 260:2315-2320(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88329734; PubMed=2843432;
 RX MEDLINE=89025644; PubMed=3178743;
 RT "Complete nucleotide sequence of the region encompassing the first
 twenty-five exons of the human pro alpha 1(I) collagen gene."
 RL Gene 67:105-115(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89025644; PubMed=3178743;
 RX Tromp G., Kuivaniemi H., Stacey A., Shikata H., Baldwin C.T.,
 RA Jaenisch R., Prockop D.J.;
 RT "Structure of a full-length cDNA clone for the prepro alpha 1(I) chain
 of human type I procollagen."
 RL Biochem. J. 253:919-922(1988).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91138770; PubMed=1995349;
 RA Maatta A., Bornstein P., Penttinen R.P.;
 RT "Highly conserved sequences in the 3'-untranslated region of the
 COL1A1 gene bind cell-specific nuclear proteins."
 RL FEBS Lett. 279:9-13(1991).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92157916; PubMed=1787829;
 RA Westerhausen A., Constantinou C.D., Pack M., Peng M.Z., Hanning C.,
 RA Olsen A.S., Prockop D.J.;
 RT "Completion of the last half of the structure of the human gene for
 the pro alpha 1 (I) chain of type I procollagen (COL1A1)."
 RL Matrix 11:375-379(1991).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98107942; PubMed=9443882;
 RA Korhonen J., Ala-Korhonen L., De Paeppe A., Nyttinck L., Earley J.,
 RA Prockop D.J.;
 RT "Analysis of the COL1A1 and COL1A2 genes by PCR amplification and
 scanning by conformation-sensitive gel electrophoresis identifies only
 COL1A1 mutations in 15 patients with osteogenesis imperfecta type I;
 identification of common sequences of null-allele mutations."
 RL Am. J. Hum. Genet. 62:98-110(1998).
 RN [7]
 RP SEQUENCE FROM N.A.

```

RA Koriko J.M., Earley J.J., Nuytink L., Depaepe A., Prockop D.J.,
RA Ala-Koriko L.;
RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF017178; AAB94054.2; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000885; Fib_collagen_C.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF01410; COLFI; 1.
DR Pfam: PF01391; Collagen; 18.
DR Pfam: PF00093; VWC; 1.
DR ProDom: PD000007; Collagen; 2.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01208; VWF_C; 1.
DR Collagen.
SQ SEQUENCE 1461 AA; 138630 MW; 9ACF6DE30EA78E21 CRC64;

Query Match
Best Local Similarity 100.0%; Score 80; DB 4; Length 1461;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCGV 15
DB 944 GTPGPGIAGGRCGV 958

RESULT 5
Q8N473 PRELIMINARY; PRT; 1464 AA.
AC Q8N473;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (Aug-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC036531; AAH36531.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000885; Fib_collagen_C.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF01410; COLFI; 1.
DR Pfam: PF01391; Collagen; 18.
DR ProDom: PD000007; Collagen; 2.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01208; VWF_C; 1.
DR Hypothetical protein: Collagen.
SQ SEQUENCE 1464 AA; 139011 MW; B0581F8D1C89DDE8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 80; DB 4; Length 1464;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCGV 15
DB 947 GTPGPGIAGGRCGV 961

RESULT 6
Q93485 PRELIMINARY; PRT; 809 AA.
AC Q93485;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

```

```

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Alpha 1 type I collagen (Fragment).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_Taxid=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RA Saito M., Kunisaki N., Hirono I., Aoki T., Ishida M., Urano N.,
RA Kimura S.;
RT "Partial characterization of cDNA clones encoding the three distinct
RT pro alpha chains of type I collagen from rainbow trout.";
RL Fisheries Sci. 64:780-786(1998).
DR EMBL: AB008373; BAA3380.1; -.
DR InterPro: IPR000885; Fib_collagen.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF01410; COLFI; 1.
DR Pfam: PF01391; Collagen; 9.
DR ProDom: PD000007; Collagen; 1.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR Collagen.
FT NON_TER
SQ SEQUENCE 809 AA; 78164 MW; 68C056A7640FCA81 CRC64;

Query Match
Best Local Similarity 93.8%; Score 75; DB 13; Length 809;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTPGPGIAGGRCGV 15
DB 292 GTPGPGIAGGRCGV 306

RESULT 7
Q910C0 PRELIMINARY; PRT; 1449 AA.
AC Q910C0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Collagen a1(I).
GN COL1A1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_Taxid=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21257802; PubMed=11358497;
RA Saito M., Takenouchi Y., Kunisaki N., Kimura S.;
RT "Complete primary structure of rainbow trout type I collagen
RT consisting of a1(I)2(1)3(I) heterotrimer.",
RT Eur. J. Biochem. 268:2817-2827(2001).
DR EMBL: AB052835; BAB55661.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000885; Fib_collagen_C.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF01410; COLFI; 1.
DR Pfam: PF01391; Collagen; 18.
DR Pfam: PF00093; VWC; 1.
DR ProDom: PD000007; Collagen; 2.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01208; VWF_C; 1.
DR Collagen.
SQ SEQUENCE 1449 AA; 137117 MW; 62EEF8A7BFD652B8 CRC64;

Query Match
Best Local Similarity 93.8%; Score 75; DB 13; Length 1449;

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC030913; AAB30913.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000885; Fib_collagen_C.
DR Pfam: PF01410; COLF1; 1.
DR Pfam: PF01391; Collagen; 9.
DR ProDom: PD000007; Collagen; 2.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLF1; 1.
DR Collagen.
KW NON TER.
FT
SQ SEQUENCE 826 AA; 80124 MW; B5BD721772BD4F24 CRC64;

Query Match
Best Local Similarity 88.8%; Score 71; DB 11; Length 826;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTPGGGAGGRCYV 15
ID 08CEF7 PRELIMINARY; PRT; 886 AA.
AC 08CEF7;
DT 01-MAR-2003 (TReMBLrel. 23, created)
DT 01-MAR-2003 (TReMBLrel. 23, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Procollagen.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL: AK028295; BAC25865.1; -.
DR SEQUENCE 886 AA; 85536 MW; 47A70AADBBF4F45 CRC64;

Query Match
Best Local Similarity 88.8%; Score 71; DB 11; Length 886;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTPGGGAGGRCYV 15
ID 014046 PRELIMINARY; PRT; 1160 AA.
AC 014046;
DT 01-NOV-1996 (TReMBLrel. 01, created)
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, last annotation update)
DE COL2A1 protein precursor (Fragment).
CN Homo sapiens (Human).
OS

```

